

Wake-up call for Japan's genomics

A US company's plans to sequence the rice genome in weeks rather than decades should inspire a more venturesome spirit in the Japanese research community.

Last week's news of a newly launched private initiative in the United States to sequence the entire rice genome in only six weeks has sent shock waves through the plant genomics community. It has prompted Japanese researchers engaged in the international Rice Genome Sequencing Project to consider a significant revision of their ten-year initiative planned for completion in 2008.

No doubt those involved in the publicly funded effort are experiencing déjà vu. The private sequencing project, to be carried out by Celera Genomics, the company established by the genome researcher J. Craig Venter, has created a situation very similar to that last year, when he announced plans to sequence the complete human genome, in direct competition with the efforts of researchers at the National Institutes of Health, the Wellcome Trust and elsewhere. Japanese researchers are now having to come to terms with what they consider is a major challenge to their own rice genome project.

Venter plans to create a commercial database containing sequence data of the 430-megabase genome (see page 545). Despite general scepticism over Celera's proposed six-week timescale, the effect of the news has been instantaneous — in a matter of days, Japanese researchers had drawn up plans to accelerate their sequencing effort, including a request for increased funding to purchase DNA sequencers. Researchers realize the urgent need to reaffirm and speed up the publicly funded project, which has so far taken a sluggish path despite rapidly intensifying competition in plant genome research from the private sector. Limited sequencing ability caused by lack of funds is partly to blame for the slow progress of the sequencing effort. Researchers also blame the government for being slow to acknowledge the importance of genome research — as a result, Japan has made a slow start in this sphere of activity, while the United States, prompted by scientists at the National Institutes of Health and

elsewhere (including Venter), enthusiastically embraced such research as a national strategy.

The challenges arising from Venter's initiative will also be an opportunity for the Japanese government to revise its policy on genome-related research. After years of urgent calls from Japanese researchers to promote genome research more energetically on a national level, the government will introduce a new programme this year to promote the commercial application of biotechnology research, particularly in genomics, by increasing funds for basic research and creating public databases of genomic information. But this is scratching the surface by international standards. Japan is far behind the United States and Europe in genome research, and urgently needs to boost not only sequencing capacity and other basic research but also the environment in which research and commerce can jointly thrive. Venture capital needs more encouragement, but the government should also promote rather than, as at present, discourage collaboration between public and private sectors. Japanese researchers themselves would do well to view commercial collaboration more positively than they do now.

Venter's approach will give him a lead in identifying sequences with technological potential. Whether that will in turn result in an appropriate balance between the right to profit from such information and the needs of the world's population is quite another matter, which now requires discussion all the more urgently. But there is nothing to stop publicly funded genome sequencing projects from making their accumulating sequences available to all. Given the obvious commercial and political interests in the sequence for rice, which is the principal food of half of the world's population, the quicker this happens, the better. Rice genome partners — Japan, the United States, the European Union, China and South Korea — please note. □

Xenotransplant caution continues

A ban on clinical trials of transplants from primates shows good judgement. Now we need data.

It's official. Xenotransplants from non-human primates would expose "recipients, their close contacts, and the public at large... to significant infectious disease risk" (see page 549). This strong language comes from the US Food and Drug Administration (FDA), which just three years ago wanted to leave the regulation of xenotransplantation to local institutional review boards. In private, some FDA officials now also argue that the question of whether to proceed with human trials of living parts of any animals is so important that a federal policy ruling is needed.

That it has taken the FDA until now to decide the obvious says much about the reluctance in the United States to regulate if this can be at all avoided. The potential benefits of using primates are tiny — breeding enough clean monkeys or baboons to substantially shorten organ waiting lists is impracticable. The potential cost is the creation of more AIDS, Ebolas or Marburgs.

It would be a mistake to conclude that the FDA's exclusion of primate donors means that other animals, such as pigs, are safe. The benefit side of the equation is better with pigs, but the risk side is largely unknown. The US Department of Health and Human Services is setting up a federal advisory committee on xenotransplantation, to include experts, representatives of other stakeholders and the inevitable bioethicists. The US administration should put human trials of pig organs on hold at least until this committee has had its say.

But it would be unfortunate were the prospect of a moratorium to have the perverse effect of deterring researchers and funders from supporting continued research on xenotransplants outside the clinic. Indeed, perhaps a rule needs to be invented, which says that whenever a government calls a moratorium, it should also make sufficient funds available for research to fill the gaps in knowledge that made it necessary in the first place. □



US firm's bid to sequence rice genome causes stir in Japan...

[TOKYO] The Japanese government's Rice Genome Sequencing Project has been thrown into turmoil by the news that Celera Genomics, the US company set up by geneticist J. Craig Venter, plans to sequence the entire rice genome in just six weeks.

Celera plans to create a commercial database of the rice genome. It would be made available to companies for US\$30 million on a five-year contract.

As reported last week in the Japanese biotechnology newsletter *Nikkei Biotechnology*, Venter plans to sequence the 430-megabase rice genome using the 'shotgun' technique. This divides the genome into small, random fragments that are then put together to produce the whole sequence.

Venter confirms that Celera has begun sequencing the rice genome, and says that it is "an important basis for future work on other plant species". He adds: "We are aware of the sensitivity over the work, given the large economic and political interest involved, but we are proceeding in order to create a database available for others."

But Venter's initiative has upset researchers involved in the international rice genome project, a ten-year initiative costing US\$200 million that plans to complete the sequencing by 2008. The project, led by Japan's Ministry of Agriculture, Forestry and Fisheries (MAFF), had aimed to have sequenced 40 per cent of the genome by 2003.

Takuji Sasaki, leader of MAFF's Rice Genome Research Program, says they will have to speed up their sequencing work. "Although we have been aware of Celera's interest in the rice genome, we are furious at this news," says Sasaki. He says the current project will adopt a new strategy aimed at creating physical maps of all 12 chromosomes by the end of the year. This will be done by mapping 'expressed sequence tags' in rice onto clones of yeast artificial chromosome, focusing on the 'gene-rich' portion of the sequence.

But he admits that Japanese researchers will be unable to compete with Celera in terms of the speed of sequencing. The US company, which is partly owned by the laboratory equipment manufacturer Perkin-Elmer, owns nearly 300 high-speed automated DNA sequencers. "We have applied for additional funding to buy DNA sequencers, but will have to wait two years for the budget to come through," says Sasaki.

Some researchers are sceptical that Celera can complete the sequencing in six weeks. "Six weeks is an exaggeration — a reliable sequence for the rice genome cannot be



Slow business: Critics say that Japan has not responded to competition in genome research.

obtained by such a fast-track method," says Michio Oishi, director of the Kazusa DNA Research Centre, Japan's first institute dedicated to the sequencing and analysis of DNA.

Oishi predicts that Venter's approach will produce "80 to 90 per cent of the sequence data", but he notes that the objectives of the two sequencing efforts are very different.

Sasaki says that while the public project aims to provide researchers with complete and reliable sequence data, Venter's effort centres around the aim of "licensing whatever he can achieve".

But he points out that the sequencing will require mapping data if the isolated genomic

sequences are to be reassembled in the correct order. Venter says he has "sufficient information to complete the sequencing work", although he adds that "Celera would be happy to partner with groups in Japan to help understand the rice genome".

"Although it would be ideal to collaborate with Celera on this project, there will always be a problem about release of the data," says Sasaki. Like the Human Genome Project, the rice project has agreed on the immediate release of the sequence data; but others would like the option of seeking patent protection on some of the sequence information.

Oishi, one of a group of leading researchers who released a statement last year urging the government to tackle genome research on a national basis, admits that "Japan has been particularly slow at responding to the intensifying competition in genome research".

Answering claims that his initiative would damage the 11-nation international rice genome project, whose members include the United States, the European Union, India and China, Venter argues that the public project is still at a very early stage and has made little progress.

"We have the set-up and the technology to proceed with the sequencing work, and although we do not mean to compete with the public initiative, we can't wait until they get their act together," he says. "Sequencing the genome is just the beginning. We must not forget that post-sequencing work — the analysis of the sequence data — requires far more work."

Asako Saegusa

...as consortium plans free SNP map of human genome

[LONDON] Ten of the world's leading pharmaceutical companies, together with Britain's Wellcome Trust, are due to confirm today (15 April) that they will fund a US\$45-million map of the variations in the human genetic code that could be linked to common diseases.

The map will not be patented and will be given free to the world's research community. It will be made available to researchers as it is being constructed, and is expected to be completed within two years.

The project aims to identify up to 300,000 variations in the 3 billion nucleotide base pairs that make up the human genome. These variations are called single-nucleotide

polymorphisms, or SNPs; one SNP is estimated to occur roughly every 1,000 bases.

In the long term it is hoped that the map will lead to more effective medication, such as 'personalized' drugs, and to the development of smaller and cheaper drug trials.

In the short term, however, an SNP map is expected to help to reduce the incidence of side effects from existing medicines, and to make the search for genes associated with diseases much faster and cheaper.

For example, by comparing SNP maps of individuals suffering from a disorder such as coronary heart disease with maps of healthy people, researchers hope to identify regions of the human genome where genes associat-

ed with the disease might be found.

The Wellcome Trust will contribute \$30 million to the project, and has already funded three pilot studies. The SNP Consortium of pharmaceutical companies will contribute the remaining \$15 million.

The consortium consists of AstraZeneca, Bayer, Bristol-Myers Squibb, Hoffmann-La



Bentley: maps are 'pre-competitive research'.

Roche, Glaxo Wellcome, Hoechst Marion Roussel, Novartis, Pfizer, Searle and SmithKline Beecham. The identification and analysis of SNPs will be done at centres involved in the international Human Genome Project.

The US National Institutes of Health has already given \$30 million in grants over three years for *ad hoc* SNP-related projects. The idea that pharmaceutical companies could fund a more comprehensive map emerged from discussions two years ago at the US National Human Genome Research Institute, says Alan Williamson, formerly vice-president for basic research at Merck and a member of the institute's governing council.

Williamson says that pharmaceutical companies were initially approached with the idea of forming a profit-making consortium, which would charge for the use of the map. But those companies that agreed to the idea surprised most observers by insisting that the data should be made available free. The Wellcome Trust's involvement was also conditional on full and free access to the data.

David Bentley, of the Wellcome Trust's Sanger Centre for gene sequencing, says consortium members regard the project as 'pre-competitive research', and have agreed not to see any SNP data before they become public.

Glaxo Wellcome is already using map data

[LONDON] The UK-based drugs company Glaxo Wellcome — which says that it has patented the idea of a single-nucleotide polymorphism (SNP) map of the human genome (see page 545) — has been attempting for some time to incorporate information about variations in gene sequences into its drug discovery programme.

The company became convinced of the feasibility of mapping the variations in nucleotide base pairs in 1997. Within four months, it had used an SNP map from late-

onset Alzheimer's patients to narrow down a possible genetic basis for the disease to just two genes — *ApoC1* and *ApoE*.

Over the past two years, the company has launched studies of patients with asthma and coronary heart disease, disorders whose genetic basis is less well known. In addition, it asks patients on clinical trials for many of its drugs for permission to include their genetic profiles in a future SNP database.

According to Allen Roses,

a vice-president of Glaxo Wellcome and worldwide director of genetics, a more complete SNP map could establish a genetic basis for handling side effects of drugs.

Roses says doctors are reluctant to prescribe an anticonvulsant drug, Lamictal, because five per cent of patients develop a potentially fatal skin rash as a side effect. If a genetic basis for such side effects were found, patients could be screened, and the drug given to the majority not predisposed to developing the rash. **E.M.**

But the consortium's members will receive an additional benefit. By helping to ensure that SNP data are freely available, the larger pharmaceutical companies will not be restricted to the many proprietary SNP databases already being developed by smaller genomics companies. Indeed, these companies are absent from the SNP Consortium, which is made up exclusively of established pharmaceutical companies.

Celera Genomics near Washington DC is compiling a proprietary SNP database in parallel with its efforts to sequence the human genome. Its president, J. Craig Venter, says he welcomes the new SNP consortium.

Venter says he considers it to be more of a "public education programme" in SNPs than a threat to companies such as Celera. He adds that he expects considerable demand for Celera's product, which, he says, is the subject of a possible patent application.

Two pharmaceutical companies outside

the consortium — Amgen and Pharmacia UpJohn — have already signed up, says Venter, adding that Celera expects to obtain between 20 and 30 million SNPs within the next 18 months, two orders of magnitude more than the SNP Consortium. Celera's database, he envisages, could also be augmented by data from the SNP Consortium.

But Allen Roses, a vice-president of Glaxo Wellcome and worldwide director of genetics, says 200,000 SNPs are enough for most practical purposes, such as clinical trials.

Roses claims that it is too late to patent an SNP map, as the concept is no longer novel. He says that the basic idea has already been patented by Glaxo Wellcome (see above) following a mapping exercise to find genes associated with Alzheimer's disease.

Furthermore, says Roses, companies marketing SNP databases will not be allowed to sell data obtained from the SNP Consortium back to a consortium member. **Ehsan Masood**

French research would benefit from less paperwork, says report

[PARIS] France should create a structure to coordinate administrative tasks across its research agencies. This is one of the main conclusions of a report on research bureaucracy, commissioned by the ministry of national education, research and technology and released last week.

The report was prepared by Marc Goujon, a researcher at the Centre National de la Recherche Scientifique (CNRS), and Gérard Chastagnaret, head of the Mediterranean Centre for the Sciences of Man, following extensive consultation in the research community.

Hardly surprisingly, the report confirms that researchers spend increasing amounts of time on administrative tasks, for which they are often poorly qualified. At the same time, the 26,000-strong CNRS, which is often

perceived as a bastion of democracy, emerges from the report as relatively efficient.

It dedicates just 8.5 per cent of its staff and 7.6 per cent of its budget to administration — a share that would be the envy of many private companies. In contrast, the computing agency INRIA allocates 35 per cent of its staff and 18 per cent of its budget to management.

The report deplores the way scientists often have to deal with five layers of regulations: Europe, the state, the regions, the agencies to which they belong and the universities where their laboratories are based. It upholds the CNRS's decentralization of administration away from Paris, through the creation of a series of regional delegations, as a model for other agencies to simplify procedures.

The report criticizes the fact that each agency tends to operate in isolation from the others, resulting in incoherence and duplication. Each agency has developed its own expensive solutions to address the Millennium Bug computer problem, for example, and each has its own scientific press. An interagency structure could centralize many tasks such as training and management of laboratory budgets, says the report.

Claude Allègre, the science minister and an Internet enthusiast, is likely to be pleased by one of the report's recommendations: that the paper mountain associated with administering research should be slashed by implementing security systems to give e-mails the same legal status as signed paper documents. **Declan Butler**

Plan to merge optical societies under fire

[SAN DIEGO] A bid by the leaders of two major US optical science societies to merge their organizations has spawned a vociferous backlash. Dissident scientists are concerned that they could be steamrolled by engineers in the proposed new organization.

Leaders of the Optical Society of America (OSA), which is made up largely of scientists, and of the International Society for Optical Engineers (SPIE), a principally engineering group, are planning to ask the members of their organizations to vote later this year on the proposed merger.

Their goal is to form an international organization that is "intended to be the primary professional society for optical scientists and engineers so as to unify the field", says OSA president Anthony Siegman, an electrical engineer professor emeritus at Stanford University.

SPIE president Paul Schenker, an electrical engineer at California Institute of Technology, says the idea of unification is "grounded soundly" and argues that it "would provide better service to the optics discipline".

But the speed and manner in which the organizations' boards have moved ahead with the proposal over the past year has prompted a number of scientists to mount a challenge in order to either slow down the process or halt it completely.

A petition drive has been started, seeking a special meeting of OSA members at a conference next month in Baltimore, where OSA critics of the merger hope to muster votes to prevent it taking place immediately.

"It is being done too hastily [and] without sufficient safeguards," says Daniel James, a quantum computer scientist at the Los Alamos National Laboratory. "There may be a need for a change in the relationship of the two societies, but it should be evolutionary rather than revolutionary."

Objections to the unification plan appear to be rooted in a cultural clash between scientists and engineers. If the merger goes ahead as proposed, the critics fear that the purely scientific endeavours of OSA members will be dominated by the applied interests of engineers.

They are also worried that the amount of bureaucracy would increase, and that an end to the traditional competition between the two groups — which they argue has stimulated innovative thinking — will reduce the number of new scientific discoveries.

Some long-standing OSA members, such as Emil Wolf, Wilson professor of optical physics at the University of Rochester in New York, also feel that the leadership has undertaken an expensive merger process without sufficient support from the membership.

"It's like a business merger — a hostile merger done in an highly undemocratic

Quantitative comparisons of SPIE and OSA

	SPIE	OSA
No. of members	14,000	12,500
No. of corporate members	320	110
Dues	\$95 US, \$105 non-US	\$80
Typical member characteristics	Interested in engineering and applications	Middle-age white male, PhD in physics and interested in applied research

manner," says Wolf, one of OSA's elite honorary members. Other critics say the OSA board voted precipitously to spend up to \$400,000 on due diligence for the merger proposal, money that may be wasted if the effort is defeated.

But Schenker argues that the effort is "eminently democratic", pointing out that an opinion poll revealed that most members of both societies supported the merger. He also estimates that SPIE has so far spent only "a couple hundred thousand dollars" on unification efforts.

OSA, which is based in Washington DC, has about 12,500 members and an annual operating budget of about \$18 million, with \$14.5 million in annual revenues coming from the publication of journals and confer-

ence sponsorship. Officials say it also has about \$29 million in assets.

SPIE, based in Bellingham, WA, has about 14,000 members and about \$15 million in assets. According to Schenker, SPIE's annual operating budget is about \$16 million.

The first test of strength between those promoting the merger and their critics will take place at the Conference on Lasers & Electro-Optics/Quantum Electronics & Laser Science Conference, which opens on 23 May in Baltimore.

The OSA dissidents want to slow down the process so it does not go to a membership vote at the annual meeting in September. They also seek safeguards to protect scientist members and more control of the assets of any unified organization.

Rex Dalton

Weapons labs tighten computer security

[WASHINGTON] Classified computer operations at the US Department of Energy's nuclear weapons laboratories are expected to return to normal this week, after a week-long 'stand-down' intended to focus employees' attention on security issues.

The 'stand-down', which began on 2 April, was ordered by energy secretary Bill Richardson. It was supposed to make each of the three laboratories — Los Alamos and Sandia in New Mexico, and Lawrence Livermore in California — suspend their normal classified operations and draw up new security plans. Operations would then recommence with the new procedures in place to prevent any possible transfer of classified computer files or other data onto non-classified computer systems.

The laboratories' computer systems remained shut down early this week as they awaited word from the department that security measures had been sufficiently enhanced. According to a spokesman for Los Alamos, the laboratories have now submitted their plans. These focus on isolating the separate networks of computers that handle classified and non-classified work. There are concerns that it is too easy to transfer files from the classified network to the unclassified one, creating the potential for classified data to be leaked by e-mail.

The energy department has previously ordered stand-downs in response to safety

or environmental problems, most recently at Los Alamos and at Brookhaven laboratory in New York state. Although irritating for staff, they are intended to convince critics that a given problem — security, in this case — is being taken seriously.

Ernie Moniz, under-secretary for energy, says he is confident that morale at the laboratories will be restored when the security clampdown is over. "There is a lot of uncertainty at the labs," he admits, adding that things will improve when the new security arrangements are implemented.

Moniz does not believe that the turmoil at the laboratories will hamper recruitment plans. He concedes that the proposed introduction of routine polygraph testing "is probably the most sensitive issue", but says it will be targeted at people with access to certain types of sensitive information.

Asked whether the department has any data to support the usefulness of polygraph testing, Moniz says: "The FBI thinks it is an important tool for its investigations."

Moniz hopes to persuade people in Washington that the laboratories already have stringent security systems. "We need to get out the message of how our security systems actually work, and that we have appropriately graded security structures. People who haven't been to the laboratories sometimes think they operate like university campuses."

Colin Macilwain

Canadian Light Source is ready to glow as final grant comes in

[MONTREAL] A grant from the Canada Foundation for Innovation (CFI) has provided the final funding needed to start building a Can\$173.5 million (US\$ million) national synchrotron at the University of Saskatchewan. The machine will be called the Canadian Light Source (CLS).

According to Dennis Skopik, director of the university's Saskatchewan Accelerator Laboratory (SAL), this is the largest investment in scientific infrastructure in Canada within a generation.

The CLS will consist of an electron accelerator and storage ring that will emit electromagnetic radiation ranging from wavelengths in the infrared up to hard X-rays, providing a research tool for a wide range of disciplines. "That's why it is a very useful tool," says Skopik. "You dial in the wavelength you want and you've got it with more intensity than with any other synchrotron facility in the world."

It will be a third-generation synchrotron, of which only a handful currently exist. The decision to fund it was taken after major facilities in Europe and the United States reportedly threatened to restrict Canadians' access unless Canada developed its own.

Skopik says the decision has to some extent corrected the situation of science funding of several years ago, when major cuts were made. But he emphasizes that there was fierce competition for the funds.

"It's an extremely good step forward for Canada," he says. "It touches fields from physics to earth sciences, medical applications — just about any science and engineering is going to benefit. I don't know of any facility that has this broad range of uses."

Of the cost, Can\$140.9 million will be paid in cash and Can\$32.6 million in kind, including three University of Western Ontario beam lines which will be moved from the Aladdin facility in Wisconsin.

CFI's Can\$56.4 million contribution constitutes 40 per cent of the capital costs. The remainder will be covered by other federal departments, the Saskatchewan government, the city, Saskatchewan Power Corporation and the universities of Alberta and Western Ontario. Can\$19 million will come from other provinces, universities and industry to build beam lines.

The federal government will pay about 55 per cent of the Can\$13.9 million annual operating costs through agencies such as the research-granting councils. The rest will be covered by other sources, including users' fees and the university itself. Building will begin within two months, and the facility is due to open in 2003.

David Spurgeon

South African funding body faces evaluation headache

[CAPE TOWN] South Africa's National Research Foundation (NRF) was officially established on 1 April. It creates a single funding agency covering all research except clinical medicine at the country's universities, technikon (polytechnics) and museums.

The new organization combines the role of the old Foundation for Research Development (FRD) with the grant-giving functions of the Human Sciences Research Council, which still exists to administer its own in-house research programmes.

Members of the NRF council were announced on 30 March by Ben Ngubane, the re-appointed Minister of Arts, Culture, Science and Technology. It will be chaired by psychologist M. F. Ramashala, vice-chancellor of the University of Durban-Westville.

Khotso Mokhele, the former president of the FRD, will act as interim chief executive officer until the council appoints a head, although he is widely regarded as the front-runner for the post.

Research support will be handled by four divisions of the NRF: natural science and engineering; humanities and social sciences; environmental and agricultural sciences; and health sciences. The health sciences division is relatively small, as the NRF only pays for fundamental research in this area; clinical research is covered by the Medical Research Council, which remains autonomous.

A fifth division consists of the national research facilities that were previously part of the FRD: the SA Astronomical Observatory, the National Accelerator Centre, the Hartebeeshoek Radio-Astronomy Observatory and the J. L. B. Smith Institute of Ichthyology.

One of the government's aims in creating the foundation is to open up new opportunities for collaboration between the natural and social sciences, as it is in favour of interdisciplinary research. But the NRF faces the difficult task of integrating two very different mechanisms for evaluating research proposals in these disciplines.

The FRD has had a complex evaluation mechanism. A major component has been the individual rating of scientists by peer review. This has required that all scientists wishing to apply for funding are personally evaluated every five years, in addition to applying annually for research grants for the following year.

In contrast, research in the social sciences has been based primarily on project evaluation. Social scientists do not seem enthusiastic about the FRD evaluation system, though Mokhele described it last month as "the envy of research communities worldwide".

Mokhele sees the current funding mechanisms as remaining in place until the end of next year, when the FRD's current set of research programmes end.

Michael Cherry

Britain promises cash for UK/SA joint research

[CAPE TOWN] Britain's science minister, Lord Sainsbury, announced during a visit to South Africa last week that the UK government will contribute a further £100,000 (US\$165,000) to a bilateral research fund, extending its activities for a further year.

The UK/South Africa Science and Technology Joint Research Fund, to which the South African government makes an equal contribution, was set up in 1996 to support projects in agriculture, biotechnology and biomedical and environmental science.

At a review of the fund's first three years of operation, held in Cape Town, Sainsbury applauded its success, saying that it has "reaped considerable

benefits for science in both South Africa and the UK".

Sainsbury announced that a memorandum of understanding will be signed between the United Kingdom's Royal Society and South Africa's new National Research Foundation (NRF; see above) to build research capacity at historically black institutions.

Since September 1996, the Royal Society has provided £250,000 a year to establish centres of excellence at historically black universities — in biotechnology and marine biology at the University of the Western Cape, chemistry at the University of Zululand, stock science at the University of Fort Hare and computational modelling for

materials science at the University of the North.

A collaborative scheme between the NRF and the Royal Society provides funds for exchange visits between these universities and counterpart institutions in the United Kingdom, and for replacement salaries for South African staff members upgrading their qualifications in the United Kingdom.

The historically black universities have seen drastic reductions in enrolment over the past five years (see *Nature* 398, 277; 1999). But Prins Nevhutalu, who coordinates the scheme at the NRF, feels this should not have a negative effect on attracting postgraduate students to a research centre of excellence.

M.C.

FDA warns on primate xenotransplants

[PARIS] The US Food and Drug Administration (FDA) has announced a *de facto* ban on clinical trials of xenotransplants — transplants of living cells, tissues and organs — from non-human primates to humans.

"Recipients, their close contacts, and the public at large would be exposed to significant infectious disease risk," the FDA warned last week in an unusually strong statement.

The step marks formal acknowledgement by the FDA of widespread concern over the risk that animal viruses might jump to recipients and then spread to other humans, causing man-made pandemics (*Nature* 391, 320–325; 1998).

Phil Noguchi, director of the FDA's division of cellular and gene therapies, says that the move was prompted by a "groundswell" of public and scientific opinion over the past two years that "non-human primates are a potential hazard".

Primates seem attractive as donors, because their tissues and organs are more like those of humans than are those from more phylogenetically distant animals such as pigs, and are therefore less susceptible to rejection. But primates are widely considered unsuitable as a source of xenotransplants (*Nature* 392, 11–12; 1998).

One reason is that the infectious disease risk is widely considered to be unacceptably high. Ebola and Marburg monkey viruses have caused large disease outbreaks in humans, while there is compelling evidence that HIV originated from monkey retroviruses. Moreover, it would be impractical to breed the large numbers of 'clean' animals that would be needed to make any realistic impact on the organ shortage.

The new FDA "guideline to industry" (see <http://www.fda.gov/cber/gdlns/xenoprim.txt>) does not explicitly forbid xenotransplants from non-human primates. But in practice the guideline does amount to a ban — it warns clinical trial sponsors that there is no point applying to FDA for approval of protocols involving such xenotransplants, since they will not be able to satisfy the FDA that it is safe.

"It took a while to craft the language," says Noguchi. The FDA feels that its remit is to rule on individual submitted trial protocols, he comments. It considers that decisions concerning a formal ban or moratorium on the technology should be taken by the US Department of Health and Human Services. The department is already in the process of setting up a federal xenotransplantation advisory committee, along the lines of the National Institutes of Health's Recombinant DNA Advisory Committee.

The FDA move has been widely welcomed. "I applaud the FDA for taking a leadership role here," says Jon Allan, a virologist



at the Southwest Foundation for Biomedical Research in San Antonio, Texas, and a member of the FDA advisory subcommittee on xenotransplantation. He says it will send a strong message internationally about the unsuitability of primates for this purpose.

But Alix Fano, of the Campaign for Responsible Transplantation, is one of several critics who say that "an out-and-out ban would have been preferable". Critics also argue that the ban sends the implicit message that species other than non-human primates — such as pigs, the current donor of choice — are therefore safe.

Fritz Bach, a xenotransplant scientist at Harvard Medical School, complains that the FDA position is confusing. "It's not unlikely that non-human primates are more dangerous in terms of disease transmission than pigs," he says. "However, we know nothing about how dangerous it would be to use pigs. Given that, how can the FDA now suggest a moratorium on non-human primate donors but maintain a policy potentially allowing pigs as donors?"

Paul Herrling, the research director of

Novartis, says that the FDA statement is "fully in line with company thinking; we think pigs are safer".

But the FDA does not necessarily support the use of pigs. "On a scientific basis, so far the dangers of non-human primates have been better demonstrated," says Noguchi. "But that does not mean that other species will or will not be a danger."

Indeed, this month Malaysia decided to slaughter 1.3 million pigs to try to contain an outbreak of Hendra virus, a disease that has jumped from pigs to humans and has already killed more than 70 people (see <http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/00056866.htm>).

"In terms of virology there is evidence that the most damaging viruses can be those that cross widely disparate [species] barriers," adds Louisa Chapman, a xenotransplantation expert at the US Centers for Disease Control (CDC). "CDC is therefore very reluctant to sign on to any argument for saying primates are more dangerous than pigs."

At the same time, she believes that the tight husbandry rules envisaged in Public Health Service guidelines on xenotransplantation — which are scheduled for release this summer — would prevent such a virus from being a problem in trials in the United States.

Noguchi cautions, however, that the Hendra virus outbreak "once again reminds us that every time we use living animals' cells, tissues or organs there is a chance of something very bad happening".

Several clinical trials of pig cell transplants are under way in the United States. But the Hendra virus will cause the FDA to think carefully, says Noguchi. "If we think that xenotransplantation using pigs is not safe we will stop it," he says. "Meanwhile we want to move ahead carefully."

Declan Butler

Serb scientists call for help to end bombing

[MUNICH] Serbian scientists have called on colleagues in NATO countries to oppose the bombing campaign in Serbia and Kosovo.

The scientists have sent e-mail messages to academic research institutions around the world. They point out that many scientists resisted the Milosevic government's attack on the independence of Serbian universities last year, and helped to force his government to admit defeat in the 1996 elections won by the opposition (see *Nature* 394, 715; 1998). The election climbdown was achieved through "three bitter winter months of peaceful demonstrations in rain and snow".

Scientists at the Institute of Molecular Genetics and Genetic Engineering in Belgrade say the fact that NATO is bombing the cities ruled by the opposition "has led to

the common feeling that NATO is against all people in Yugoslavia".

Scientists at the institute appear not to have been worn down completely by the international sanctions imposed against Serbia in 1992.

Despite the difficulties of importing chemicals, equipment and journals, and the absence of international collaboration, they have been able to keep up some research. Many have also had work published in international journals. The researchers have even managed to complete a new building, which was opened last October.

But now, they say, they travel to work in fear that they will find their laboratory destroyed by bombs, and travel home in fear of being hit themselves.

Alison Abbott

German centre risks women-only job ads

[MUNICH] One of Germany's largest national research centres, the Research Centre Jülich, has advertised tenure-track positions reserved exclusively for women in a bid to increase the number of women occupying top research positions.

The move has been widely applauded in principle. But it has also raised some eyebrows in other research institutions. Personnel officers say they had been led to believe that advertising women-only positions for scientific posts is illegal because it discriminates against men.

But Sybille Krummacher, a scientist at the Jülich centre, which conducts research into energy and the environment, says: "It was time to stop the continued analysis of why women are under-represented in top scientific positions, and start taking bold action."

She claims to have identified a loophole in the legal situation which forbids sex discrimination in jobs where the sex of an applicant is irrelevant. A federal law forbidding discrimination in job advertising was upheld by the European Court in 1995 in a ruling on a case in which the city of Bremen wanted to advertise jobs specifically for women. But the court also said that employers could take steps to promote the "realization of equality of opportunity".

Women-only tenure-track positions fall

into this category, argues Krummacher, as they keep highly trained women in research at exactly the time when many are forced to choose between starting a family and keeping a career. The high drop-out rate at this critical time means that women do not have equal opportunities for senior positions, she adds.

A recent report by Germany's Science Council, the Wissenschaftsrat, points out that, although half as many women as men complete PhDs, only 4.5 per cent of top academic jobs are occupied by women (see *Nature* 393, 402; 1998).

This difference is blamed on Germany's hierarchical academic system. This keeps young researchers as juniors dependent on their professor or laboratory chief, which makes it difficult for any scientist to establish a secure research career before the age of 40. Combined with a lack of child-care facilities, this discourages women who do not want to sacrifice their chance to have children.

Jülich's tenure track programme for female scientists, which offers three positions per year for the next three years, addresses this problem by advertising explicitly for women who do not necessarily wish to "abandon the idea of having a family".

The Max Planck Society (MPS), which runs 82 research institutes, has a programme to try to increase the number of women in

top positions. Only five of its current 230 directors are women. The MPS has avoided possible legal conflicts over discriminating against men by using its own funds to support nine five-year 'C3 positions', equivalent to the lower rank of professor, and by calling for nominations from its eighty-odd research institutes, rather than advertising.

A spokesman for the MPS calls the decision of the Jülich Research Centre "very courageous", but fears that its legal basis may be shaky. Germany is a very male-dominated



Förster: 'a step in the right direction'.

society, he says, and it cannot be ruled out that some male researcher will bring a discrimination case against Jülich.

Detlev Ganten, president of the Helmholtz Society, the umbrella organization for Germany's 16 national research centres (including Jülich),

says that there is no general support within the society for positions earmarked for women. Aside from the legal problems, he says, many women who are successful in science feel that there is a stigma attached to such posts.

The Helmholtz Society has recently introduced a number of other measures, he says, such as recommending stronger representation of women in recruitment selection committees, and trying to improve child-care facilities. It is promoting consciousness-raising at all levels, from special mentoring programmes with school students to the creation of a working group to advise the society's own management.

But Andrea Förster, a researcher at the Geology Research Centre (GFZ) in Potsdam — also a national research centre — describes the Jülich tenure-track positions as "a step forward in the right direction". She points out that, while women make up a third of the GFZ's scientific staff on short-term contracts, only six of the 280 permanent scientists are women.

Förster, one of the young east German scientists kept on at the institute when it was re-founded after reunification in 1990, does not think that improved child-care facilities alone will help improve the lot of women in science. She points out that there was a high proportion of women in science during the communist era, when child-care facilities were taken for granted, but there were still not many in high positions.

Förster blames the structure of science in Germany, and in particular the time it takes for a scientist to get on to a career track. The Jülich programme is a direct challenge to this structure, she says, which goes straight to the heart of the problem.

Alison Abbott

'Timeline' celebrates a century of physics



[WASHINGTON] An illustrated 'timeline' tracking the past 100 years in the history of physics was unveiled in Washington this week by the American Physical Society (APS). The series of posters was created to celebrate the society's centenary, and features more than 90 Nobel laureates.

The 11 posters form a 23-foot mural (above), known as 'A Century of Physics'. Copies will be sent to every high school and college in the United States. A hypertext

version of the timeline is also available at www.timeline.aps.org on the web.

It was unveiled at a reception that also celebrated the award of an APS fellowship to Representative Rush Holt, former assistant director of the Princeton Plasma Physics Laboratory (see *Nature* 394, 411; 1998). Its sponsors include the US Department of Energy, the Richard Lounsbery Foundation, IBM, Lucent Technologies, National Science Foundation and United Parcel Service.

NIH stem-cell guidelines face stormy ride

[WASHINGTON] The US National Institutes of Health (NIH) moved a step closer to funding stem-cell research last week. A working group met to refine draft ethical guidelines that scientists would have to obey to meet congressional criteria for the protection of embryos.

The working group deliberated over issues of informed consent, and wrestled with obscure but politically important definitions of words such as 'pluripotent'. But opponents of the research who attended the meeting lambasted the NIH for proceeding on a path that, they claim, contravenes the spirit and the letter of a law that bans federally funded embryo research.

Maggie Wynne, dispatched by the Pro Life Caucus of the House of Representatives, urged NIH director Harold Varmus to reverse his decision to fund the research.

The existing law prohibits funding for research in which embryos are "destroyed [or] discarded", and extracting stem cells from embryos requires their destruction. The Department of Health and Human Services interpreted the law in January as saying that federal funds may finance work on existing stem cells, but not their derivation from embryos, which would have to be financed privately (see *Nature* 397, 185; 1999). (Stem cells can also be derived from aborted fetuses, which is permissible under federal law.)

At the meeting, the opponents said that recent work may obviate the need to extract stem cells from embryos. This research showed that more specialized stem cells in adult tissues, from brain to bone marrow, can give rise to a variety of tissue types including nerve cells, cartilage, bone and connective tissue.

Richard Doerflinger, of the National Conference of Catholic Bishops, declared that extracting stem cells from embryos may become "irrelevant". He urged the NIH to invest instead in research on adult-derived stem cells for cell and tissue replacement.

Using adult stem cells to derive specialized cells and tissues for therapies would also answer the problem of tissue rejection inherent in use of embryonic stem cells. Provided that the science develops appropriately, an adult could simply have his or her own stem cells harvested and instructed to differentiate into a particular type of tissue; for instance, insulin-producing pancreas cells for a diabetic, or dopamine-producing neurons for a patient with Parkinson's disease.

By contrast, Brigid Hogan, a cell biologist at Vanderbilt University in Tennessee, and a member of the NIH working group, estimated that about 20 immunologically different lines of embryonic stem cells would need to be established for therapies to allow immune compatibility with most of the population.

Shirley Tilghman, a molecular biologist at Princeton University, New Jersey, and co-chair of the working group, emphasized that its mandate was to produce guidelines and not to pronounce on whether the government ought to fund stem-cell research. But, given the opposition among anti-abortion groups and their allies on Capitol Hill, the guidelines seem destined to become a political as much as an ethical document.

In the event of a showdown between stem-cell research opponents and supporters in Congress, vacillating politicians being wooed by the promise of the research could point to the guidelines as an assurance that federally funded research would be conducted ethically. But, no matter what the guidelines say, opponents are unlikely to be satisfied by them.

At the heart of the proposed guidelines is the requirement that stem cells be extracted only from embryos left over from fertility treatments. Those doing the extraction — either companies or scientists working with private funds — would have to provide to federally funded scientists receiving the cells documentation that the embryos were created for infertility treatment and not for research.

They would also have to guarantee that the woman who donated them did not experience, according to the draft, "undue or even subtle pressure to donate". She must also be informed that commercial products could be developed from the cells, and who would own the products or patents. And the draft forbids the use of stem cells for reproductive cloning and the creation of human-human or human-animal chimeras.

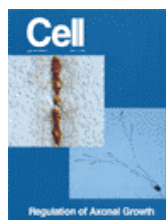
In a separate move, a document released at



Protests ahead: anti-abortionists, demonstrating in Washington, remain opposed to such research.

the NIH meeting made it appear that President Bill Clinton's National Bioethics Advisory Commission will be more permissive when it makes recommendations on stem-cell research in June. In a letter sent to Varmus by Harold Shapiro, chair of the commission, Shapiro wrote that "we seem to be coming to general agreement" that federal funding should be permitted for both the derivation and research use of stem cells derived from embryos left over from fertility treatments and from aborted fetuses. **Meredith Wadman**

Cell sale to Elsevier will boost web ventures



[PARIS] The independent biweekly journal, *Cell*, previously owned by Benjamin Lewin, has been bought by the Dutch company Elsevier Science, in a strategic move that the company says is

aimed at consolidating its electronic publishing activities in life sciences.

Lewin will remain as editor-in-chief, following the purchase of Cell Press — whose other titles are *Molecular Cell*, *Immunity* and *Neuron* — according to Geert Noorman, managing director of Elsevier Science's Life Sciences Division.

"Cell Press will stay as an independent brand with the same editorial policy," says Noorman. "Why change a winning team?"

He adds that the primary motivation behind Elsevier Science's move was the

desire to acquire a flagship journal for Elsevier's electronic publishing activities. The deal will provide Cell Press with the infrastructure to allow it to develop its electronic activities, he says.

By combining the brand names of Cell Press with the large range of Elsevier Science journals in life sciences, the company hopes to become one of the "first-stop shops" for scientific information on the web, says Noorman. "It's the battle of the bookmarks."

The company has not given any financial details of the deal, but the sum paid is widely believed to have been in excess of US\$100 million.

Cell was set up in 1974 by the Massachusetts Institute of Technology with Lewin — previously a senior editor at *Nature* — as its editor. The magazine was subsequently bought by Lewin and has been owned by him ever since. **Declan Butler**

US could face threat to its lead in immunology research, warns report

[WASHINGTON] The United States is the world leader in immunology research, says a report released last week by the National Academies of Science and Engineering and the Institute of Medicine. The country is also said to be "dominant" in the four major subfields of immunology: cellular immunology, molecular immunology, immunogenetics and clinical immunology.

But the report notes that the United States is only "among" the world leaders in some areas of these subfields, including lymphocyte development and self-recognition, tumour immunology, and transplantation and immunosuppressive drugs. And it warns that US leadership may be threatened by the growing cost of maintaining mouse facilities, and increasing competition from Europe and elsewhere.

The report gauged the US position by several measures: "reputation survey", in which experts in each subfield were asked to compile lists of the top people in their field (60 to 70 per cent of those named were US scientists); citation analysis of a "high impact" immunology database (US laboratories garnered 66 per cent of the

citations); and publication analysis in leading journals.

US-based researchers were found to have contributed about two-thirds of the immunology papers published in *Nature* between 1995 and 1997, and about three-quarters of those published in *Cell*, *Science* and *Immunity* in the same period.

Australian call for repeal of restrictive cloning laws

[SYDNEY] Two Australian authorities have come up with similar recommendations on human cloning. The Australian Academy of Science and the Human Genome Project (HUGO), which met in Brisbane last month, examined the regulation and ethics of human cloning research.

Both bodies distinguish between reproductive cloning — which the academy says should be "prohibited" and HUGO's ethics committee urges "should not be attempted" — and therapeutic cloning of stem cells, tissues and organs. Each supports therapeutic cloning under ethical guidelines, peer review and public scrutiny. The academy says "unduly restrictive legislation" in three of the six states of Australia needs to be repealed to "capture benefits from recent progress in cloning research". HUGO acknowledges that a question "remains to be

resolved" on whether all cells created by somatic cell nuclear transfer should be regarded as embryos.

Merger creates world force in biotechnology

[LONDON] Two key names in UK biotechnology — Merlin Ventures and the Rothschild Bioscience Unit — are to merge to form one of the largest teams in the world dedicated to investing in biotechnology. The new company, Merlin Bioscience, starts with £150 million (US\$242 million) in management funds, and will also run venture capital and investment management operations.

Merlin Ventures had specialized in seed and early stage investments in the United Kingdom, while Rothschild had targeted later stage companies around the world, mainly in the United States. The new company hopes to cover a broader range of investments, and to invest more in Europe.

Germany's innovation incubators underperform

[MUNICH] Innovation incubators in Germany are not efficiently meeting their aim of promoting technology transfer between start-up companies and universities,

according to a new study. The study, conducted by the University of Flensburg, surveyed all companies in the 22 incubators in the Berlin and Brandenburg area about their motives for selecting their location.

Most cited low rent and good technical infrastructure, with cooperation among neighbouring companies or university institutes as less important. Also many young companies stay for more than five years, and space is often let to well-established companies. The study recommends that the concept of technology incubators should be reconsidered, and, if necessary, that new investment should be stopped.

Survey puts grants for women in the spotlight

[LONDON] A consortium of UK funding agencies has initiated an independent national survey of academic staff to find out why women do not apply for research grants in the same numbers as men. The survey, by Social Community Planning Research, aims to identify any obstacles that deter women from applying for research grants.

Recent reports by the Wellcome Trust and the Medical Research Council found that far fewer women than expected apply for biomedical research grants (see *Nature* 390, 431; 1997).

US inquiry launched into radar patent dispute

[SAN DIEGO] The US Department of Energy (DOE) has opened an inquiry into a patent dispute on micropower impulse radar technology. An Alabama man is alleging that the Lawrence Livermore National Laboratory in California misused his discovery.

The DOE inquiry into its laboratory's actions was spurred by a congressional report released last week, raising questions about the lab's conduct in the four-year-old patent dispute. Larry Fullerton of Huntsville claims that a former Livermore scientist, Thomas McEwin, infringed Fullerton's 1987 patent on the technology when McEwin secured a patent in 1994. Livermore officials, who have licensed the technology to a number of firms, deny any impropriety.

Singapore aims to boost knowledge economy

[LONDON] The deputy prime minister of Singapore has announced a US\$1 billion Technopreneurship Investment Fund as part of the government's commitment to encouraging entrepreneurs in technology. The government also said it would review current regulations on 'technopreneurs' starting businesses, so that it could try to

remove any obstacles and build a more supportive environment and infrastructure.

The government has outlined further measures to enable Singapore to make the transition from a "capital-intensive manufacturing and services economy to a globally competitive advanced knowledge economy with a vibrant technopreneurship sector within the next decade". Measures include giving universities more autonomy to manage their operations.

Animal rights protesters wreak havoc at labs

[WASHINGTON] Members of the Animal Liberation Front ransacked a dozen research laboratories at the University of Minnesota last week. They stole 116 research animals, destroyed computers, microscopes and equipment, and disrupted experiments on Alzheimer's disease, Parkinson's disease, Huntington's disease and brain tumours, among others.

In a press release, the organization claimed "full responsibility for the strategic, nonviolent liberation" of pigeons, mice, rats and salamanders. The university said that losses are expected to total several million dollars. Local police and the Federal Bureau of Investigation are investigating, but no arrests have been made.



The full text of news items on this page – and further details and background information about the Unesco/ICSU World Conference on Science, to be held in Budapest from 26 June to 1 July – can be found through *Nature's* website at <http://www.nature.com>

Data protection and bioethics initiatives will seek backing

[LONDON] Delegates attending the World Conference on Science could be asked to endorse the creation of a “universally acceptable legal framework” that would protect access to scientific information “that takes into account the interests of all links in the scientific information chain”.

They may also be asked to support closer links between science and other systems of and approaches to knowledge “for their mutual enrichment and benefit”. And Unesco’s committee on bioethics, which has already produced a statement on protecting the human genome, could be asked to extend its activities to other genomes to prevent “exclusivity in the control of life processes”.

These are among the suggestions in a preliminary draft of the *Science Agenda — Framework for Action*, which delegates will be asked to agree to at the conference. The document is intended to indicate specific ways of putting into effect the conclusions of a parallel document, *Declaration on Science and the Use of Scientific Knowledge*, which the delegates will also be asked to adopt.

The *Agenda* is controversial. Some delegates who have seen the draft claim it is long on rhetoric but short on practical proposals, and omits any reference to, for example, the need to support particular disciplines and to encourage excellence in science. But Unesco officials point out that the draft will probably be significantly revised before the final version is submitted to the conference.

The full text of the draft *Agenda* is on <http://helix.nature.com/wcs/m00f.html>. The text of the draft of the *Declaration on Science* that will accompany it can be found on <http://helix.nature.com/wcs/m00s.html>

Full text: <http://imagine.nature.com/wcs/a21.html>

British plan to train science journalists from the South

[LONDON] Two British organizations are proposing to open an International Centre for the Communication of Science, to train science journalists, broadcasters and exhibition organizers from developing countries.

The plan is being drawn up by the Science Museum and the British Association for the Advancement of Science, which have a long tradition of communicating science to the public. The centre would be based in a building devoted to public debate on science that they are already planning to construct and occupy jointly on the museum’s land in South Kensington, London.

The proposal to add an international centre to the plans was endorsed last week by Federico Mayor, director-general of the United Nations Educational, Scientific and Cultural Organization (Unesco). At a meeting held at the museum to present plans for the World Conference on Science, Mayor told those responsible for the idea that “you can count on the full support of Unesco”.

Unesco officials say that such a centre

would fit in well with a strategy that they are keen to see endorsed at the conference. This is to identify centres of excellence in developed countries which can offer to those from developing countries intensive training courses in specific skills or areas of knowledge.

“We would very much like to see several proposals along these lines endorsed in Budapest,” says Maurizio Iaccarino, Unesco’s assistant director-general for science.

The proposal for the London centre emerged during discussions last year between Iaccarino, John Durant, assistant director for communication at the Science Museum, and Peter Briggs, executive secretary of the British Association.

“The agenda of the world conference stresses the need to set up a constructive dialogue on the ways in which science has an impact on society,” says Durant. “We feel that our idea fits the category of ‘modest and realistic proposals’ that many people argue need to be on the table in Budapest.”

Full text: <http://imagine.nature.com/wcs/a18.html>

Africans team up to boost basic research

[LONDON] Science faculties from universities in six countries in East and Southern Africa have agreed to form a network to help develop research and higher education in poorer parts of the continent.

The network will comprise initially universities in Tanzania, Kenya, Zambia, Zimbabwe, Botswana and Swaziland. The project was proposed by Mayunga Nkunya, dean of the science faculty at the University of Dar es Salaam in Tanzania. It was announced at a conference on sciences for

development held in Arusha, Tanzania.

Nkunya says many scientists see little point in doing research while their countries remain mired in poverty, military conflict and political instability. The network will give help and mentoring to staff and students in universities.

The Arusha conference promoted the idea that investment in research is a key to development, and resolved to emphasize this message at the world conference.

Full text: <http://imagine.nature.com/wcs/a11.html>

Japan’s candidate for Unesco top job promises efficiency drive

[PARIS] Japan is bidding to play a more public role in international affairs. The government has nominated Koichiro Matsuura, its ambassador to France, as a candidate for the soon-to-be-vacant post of director-general of Unesco.

Matsuura is widely considered to be one of the three front-runners for the job. He is not offering a radical reform of the role of the agency, but more efficient delivery of its existing programme, with greater emphasis on education. But he says he aims to be “realistically ambitious” in his aims and



Matsuura: Unesco needs US support.

objectives for the organization.

This approach is likely to appeal to some of the agency’s Western donors, particularly those keen to increase the cost-effectiveness with which their contributions are used.

Less certain is whether the apparent absence of a grandiose vision will inspire support from developing

countries which frequently like to see Unesco as a champion of their grievances.

In an interview with *Nature*, Matsuura said that, if elected, his priorities for Unesco would be a more efficient bureaucracy and enhanced efforts to persuade the United States to rejoin the organization. “US re-entry is very important,” says Matsuura. “We cannot hope to have meaningful discussions — for example about issues related to science — without the United States as an active participant.”

Full text: <http://helix.nature.com/wcs/c10.html>

Say no to boycott of AIDS conference

Sir—Your columns have included speculation that the International AIDS Society might countenance a boycott of the International Conference on AIDS to be held in Durban, South Africa in July 2000 (*Nature* **398**, 8; 1999). This issue has arisen because of the refusal of the South African government to routinely offer antiretroviral drugs (ARVs) to HIV-infected pregnant women, in spite of the fact that short courses of ARVs towards the end of pregnancy can reduce the likelihood of HIV transmission to infants by up to 75 per cent.

We are firmly opposed to a boycott, which would only harm the interests of those whom the well-intentioned proponents of this idea would seek to serve.

The South African government has been accused of “the murder of babies” and of having a policy that many would have considered racist, had it been implemented

by the apartheid regime. The government insists that the cost-benefit of providing ARVs to HIV-positive pregnant women should be demonstrated before any change in policy is made. But critics argue that these government studies have already been done and show that the expenses associated with the purchase of drugs, formula and clean water, and the need to care for the uninfected orphans, will be surpassed or at least equalled by the costs of providing medical treatment to the many thousands of babies who will be born with preventable HIV disease. An estimated 600,000 HIV-infected infants are born worldwide each year, 60,000 of them in South Africa.

The biennial International Conference on AIDS gains more media coverage than any other medical or scientific conference. The media barrage probably does more to promote HIV awareness and prevention

than the efforts of the scientists who present their work at the meeting. We believe that the publicity surrounding the Durban conference has the potential to translate into millions fewer new cases of HIV infection in developing countries.

Until now, South African government policy on ARVs has not received much international media scrutiny. This will change during the weeks before the conference. The government may find that a well-attended conference, with its media presence, creates more pressure to change policy on drugs for HIV-infected pregnant women than a boycott ever would have.

Mark A. Wainberg
(President)

*International AIDS Society,
McGill University AIDS Centre, Jewish General
Hospital, 3755 Chemin Côte-Ste-Catherine,
Montreal, Quebec H3T 1E2, Canada*

Take the best from both patents systems

Sir—As a patent lawyer, I read with interest your report on the European efforts to persuade the United States to switch to the first-to-file system for determining priority on patentable inventions¹. While the first-to-file system appears to be simpler, and the norm outside the United States, it would be a mistake to abandon the American system based on ‘first to invent’.

The apportionment of credit for inventions is not a mere intellectual concern because bonuses and promotions are often tied to productivity measured by patentable research. But many researchers are unable to voice their claims at the time of filing a patent because others are claiming more credit than is due to them. Reasons for such reticence include job insecurity, change in employment, immigration concerns or just the ordinary politics that accompanies all human activities.

Unlike a scientific research paper, a patent issued in the United States can be invalidated for improper naming of inventors, as measured by the conception of the invention. This includes omission and improper inclusion of inventors. Thus, the American system provides true inventors with a safeguard that can be invoked if they provide reasonable evidence of their activities. Correction of inventorship is a common hurdle faced by European patent applications filed in the United States, owing to lax standards in assigning credit

elsewhere. Interestingly, in the same issue of *Nature* Michael J. Larkin² reports that established UK molecular biologists published an average of only one paper every 3.8 years in their early careers. The US system is not perfect but it does address a problem that is mostly just debated elsewhere.

A potential compromise would be to measure priority of the patent itself by the first-to-file system but also require the naming of the inventors. Inventors would be able to appeal if their names were excluded. This would retain the threat of invalidation of tainted patents while incorporating the efficiencies of the first-to-file system.

Rattan Nath

1830 Euclid Avenue, Berwyn, Illinois 60402, USA

1. Masood, E. *Nature* **397**, 457 (1999).

2. Larkin, M. J. *Nature* **397**, 467 (1999).

Perhaps Janne Kotiaho’s concerns about mis-citation should not be limited to articles originating in non-English speaking countries (*Nature* **398**, 19; 1999). And perhaps citations should not be used uncritically for any serious purpose — at least until abstracting services become more accurate, and discover the tradition, particularly in medical journals, of giving degrees and professional qualifications after authors’ names.

Nigel L. Brown

*School of Biological Sciences, University of
Birmingham, Edgbaston, Birmingham B15 2TT, UK*

Going cold on ‘snowball Earth’ theory

Sir—Paul Hoffman and Adam Maloof¹, in discussing the paper by D. M. Williams *et al.*², misconstrue the large-obliquity hypothesis for low-latitude glaciation near sea level during the Proterozoic³ and overlook relevant geological evidence.

Seasonality would be greatly amplified globally with a large obliquity (> 54°). The solstitial climate would be severe in high latitudes but more benign near the Equator, where mean annual insolation would be minimal and where substantial snowfall could occur. The net accumulation of solstitial snow through successive equinoxes, necessary for low-latitude glaciation, would be aided by the high albedo of snow and ice. Importantly, Neoproterozoic glaciogenic deposits in Australia that formed near the palaeo-

On the trail of the prolific Dr Path

Sir—While using an electronic abstracting service I was surprised to see an additional author on one paper. This author, D. Phil, had a good publication record in a variety of fields — 76 papers over the past 18 years.

I also noted reasonable publication rates in medical research for M. R. C. Path (48), B. Chir (31) and F. R. C. Path (22). The Paths may be related as they have published jointly on four occasions. But F. R. C. Anaes (7), M. I. Biol (5) and C. Chem (4) do not seem to have been so productive.

equator⁴ contain seasonal freeze-thaw wedge structures in permafrost regolith that indicate marked seasonal changes of temperature^{3,4}.

The 'snowball Earth' hypothesis for low-latitude glaciation⁵ is difficult to simulate^{6,7} and conflicts with geological evidence. The late Neoproterozoic stratigraphic record⁸ shows no sign of the drastic lowering of sea level that would accompany global glaciation. Moreover, seasonal freeze-thaw structures could not form near the Equator on a snowball Earth because the very low global temperatures would inhibit seasonal variation⁹.

George E. Williams

*Department of Geology and Geophysics,
University of Adelaide, SA 5005, Australia*

- Hoffman, P. F. & Maloof, A. C. *Nature* **397**, 384 (1999).
- Williams, D. M. *et al.* *Nature* **396**, 453–455 (1998).
- Williams, G. E. *Earth Sci. Rev.* **34**, 1–45 (1993).
- Schmidt, P. W. & Williams, G. E. *Earth Planet. Sci. Lett.* **134**, 107–124 (1995).
- Hoffman, P. F. *et al.* *Science* **281**, 1342–1346 (1998).
- Crowley, T. J. & Baum, S. K. *J. Geophys. Res.* **98**, 16723–16732 (1993).
- Walsh, K. J. & Sellers, W. D. *Global Planet. Change* **8**, 219–230 (1993).
- Preiss, W. V. *Geol. Surv. S. Aust. Bull.* **53**, 438 (1987).
- Sellers, W. D. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **82**, 217–224 (1990).

Denmark lacks coherent policy on basic research

Sir — Recent months have seen a lively, and at times bitter, debate in the Danish media on the future direction of scientific research. Much has been made by the government of the need to shift towards more applied research to maximize the short-term benefits of public investment. However, we would suggest that more critical problems exist that must be addressed immediately to ensure the long-term health of Danish science. Chief among these are a poorly funded and misdirected policy on basic research funding, and conditions of employment that restrict the research opportunities of young scientists.

Danish science is moderately well funded¹. We have modern facilities, an excellent level of technical support and a buoyant biotechnology sector². What is sorely lacking is a coherent policy on the funding and nurturing of basic research. Entry-level appointments (assistant professor) have a heavy teaching load and no support for scientific staff. Young scientists cannot improve their situation by writing grant applications, since the funding available to the research councils allows little, if any, support for salary components. Such restrictions are making assistant professorships increasingly unattractive,

with limited long-term prospects. This situation is only alleviated by the benefaction of senior scientists and charitable foundations, and occasional directives in selected areas which allow young scientists to develop independent research.

Further obstacles exist in the recruitment process: new positions are often focused on narrow research areas and only advertised locally (in Danish). Recent well-intentioned legislative changes have not fully addressed these problems.

Such an inflexible system (which often obliges scientists to spend their entire career in the same institute) is ill-equipped to adapt to the rapid development of new areas in basic research. The only surprise is that Danish science has remained so competitive for so long. How long this will continue to be the case is unclear when there is little to attract young scientists. Without a competitive basic research component, the ability to foster novel applied research (so beloved of the present government) will be severely eroded.

Similar criticisms have been levelled at Sweden. In its case at least some of these criticisms are now being taken to heart, as can be seen by the establishment of openly competitive, well funded, junior faculty positions at the new Centre for Molecular Medicine in Umea. We hope that similar initiatives will be taken in Denmark.

Michael Ibba, Thomas Bentin

*Centre for Biomolecular Recognition,
Institute of Medical Biochemistry and Genetics,
Panum Institute, Copenhagen University,
2200 Copenhagen N, Denmark*

- May, R. M. *Science* **281**, 49–51 (1998).
- Horton, B. *Nature* **395**, 412–413 (1998).

Green revolution still too green

Sir — L. E. Drinkwater *et al.*¹ report promising isolated studies on environmental quality and productivity in crop production. But these studies should not be used to discredit high-yield crop production technologies globally.

We do not dispute David Tilman's call, in his News and Views article² accompanying the paper by Drinkwater *et al.*, for a greener agriculture based on ecological principles. But we do not think that as things stand we have the ability to cope with the probable doubling in global food demand in the twenty-first century with only organic nutrient inputs.

The organic nitrogen production schemes investigated by Drinkwater *et al.* used essentially twice the land area of their 'conventional' system to produce equivalent amounts of grain. But the best arable land is

already used for crop production. Therefore expansion of agriculture to produce the required organic nitrogen would necessitate use of marginal land or further encroachment on natural ecosystems, resulting in soil degradation and loss of biological diversity.

Additionally, agronomic technology has advanced considerably beyond the 'conventional' practices used by Drinkwater *et al.* In industrialized countries, the mould-board plough has been largely replaced by no-till and conservation tillage practices. Single applications of fertilizer are being replaced by multiple applications with amounts and timing matched to crop requirements. These practices result in improved soil conservation and quality; greater efficiency in fertilizer use; higher and more stable yields; and they also minimize the land area required for crop production.

Progress towards a 'greener' agriculture will come from continued improvements in modern high-yield crop-production methods combined with sophisticated use of both inorganic and organic nutrient sources, water, crop germplasm, pest management and beneficial organisms.

Thomas R. Sinclair*, Kenneth G. Cassman†

**USDA-ARS, Agronomy Department, University of
Florida, Gainesville, Florida 32611-0965, USA
†Department of Agronomy, University of Nebraska,
Lincoln, Nebraska 68583-0915, USA*

- Drinkwater, L. E. *et al.* *Nature* **396**, 262–265 (1998).
- Tilman, D. *Nature* **396**, 211–212 (1998).

Fertile grounds for a lively debate

Sir — I am not surprised that Hermann Bondi¹ was puzzled by my correspondence with Roger Short^{2,3}, which concerned the relative importance of fertility and agriculture in the population explosion.

I myself was puzzled by Short's response, because I was not arguing that fertility is influenced by nutrition; in fact I had been naively unaware that this is the subject of a lively controversy.

I simply meant that an increase in fertility (from whatever cause) could not by itself have resulted in population growth because the additional babies would have starved. Without the agricultural advances of this century, the world population would not now be approaching six billion, no matter what the fertility rates were.

Stephen G. Warren

*Department of Atmospheric Sciences, University of
Washington, Seattle, Washington 98195-1640, USA*

- Bondi, H. *Nature* **397**, 644 (1999).
- Warren, S. G. *Nature* **397**, 101 (1999).
- Short, R. *Nature* **397**, 101 (1999).

Cryptic clues to clock function

David Whitmore and Paolo Sassone-Corsi

The discovery that plant photopigment-like molecules are essential for the mouse biological clock adds to knowledge of the clock's molecular mechanism.

There is no point owning a watch unless you can set it to local time. The same is true for the internal circadian clock, which operates with a 24-hour periodicity in most plants and animals. Setting, or 'entraining', this biological clock requires photopigments — molecules that can detect light. Although the identity of the photopigments involved in circadian photoreception remains elusive, they are thought to be very different, both structurally and biochemically, from molecules involved in other aspects of the circadian clock. But on page 627 of this issue, van der Horst *et al.*¹ show that the mouse counterparts (or 'homologues') of plant photopigments known as cryptochromes are required for normal functioning of the mouse circadian clock. This surprising result brings putative light-sensitive molecules directly into the central clock mechanism.

The mouse cryptochromes (Cry1 and Cry2) were discovered by accident during the identification of molecules involved in the repair of DNA damaged by ultraviolet light^{2,3}. These molecules have similarities to DNA repair enzymes called photolyases, but they also contain a region similar to the blue-light-responsive molecules, or cryptochromes, found in plants. Biochemical studies showed that these cryptochromes are not involved in repairing damaged DNA^{2,3}. So what are these plant-like pigments doing in the mouse?

The most powerful way to answer this question is to block the function of cryptochrome molecules by disrupting their DNA sequences (a process known as homologous recombination). Van der Horst and colleagues¹ used this technique to generate mutant mice that lacked either functional *cry1* or *cry2* genes, or both. The most dramatic results were seen in the mice lacking both genes. When these animals were maintained on a 24-hour cycle of 12 hours of light and 12 hours of darkness, their behaviour on a running wheel was identical to that of normal mice — mice are nocturnal animals, and will run on wheels mainly during the night. But when these mutant mice were placed in constant darkness, they ran randomly at all times of the day. The mice showed this arrhythmic behaviour straight away, suggesting that their circadian clock



Figure 1 A plant–animal connection. Van der Horst *et al.*¹ have identified the mouse homologues of plant photopigments, known as cryptochromes, in the circadian clock.

had been disrupted. So, cryptochromes seem to be essential for normal clock function, and may well be involved in the central clock mechanism. Unfortunately, it is not clear how cryptochromes interact with other known mouse clock molecules, such as Clock, Period or Timeless, or how the molecular basis of the clock is affected in these mutant mice.

The authors also found that the lack of functional *cry1* alone causes the clock to run faster by over an hour. As in previous studies⁴, the lack of *cry2* has the opposite effect — the clock slows by about the same duration. These diametric effects clearly show that *cry1* and *cry2* do not overlap functionally, a concept that is supported by the different location of the two cryptochromes within the cell; whereas *cry1* is found only in the mitochondria, *cry2* is present in both mitochondria and the nucleus. The presence of cryptochromes in the mitochondria is interesting because all other putative clock molecules are involved in gene transcription and, presumably, act in the nucleus.

The new findings are very timely. Cryptochromes have already been proposed

as the circadian photopigment in a number of model systems, from the plant *Arabidopsis thaliana* to the fruitfly *Drosophila melanogaster*^{5,6}. In *Arabidopsis*, the use of specific mutants has revealed that both the red-light-sensitive photopigments, phytochromes, and the blue-light-sensitive cryptochromes are involved in setting the clock. In one interesting *Drosophila* mutant there is no oscillatory expression of the clock gene *period*, and the usual suppression by light of another clock gene, *timeless*, is blocked⁶. Characterization of this mutant led to the identification of *crybaby*, a *Drosophila* homologue of a plant cryptochrome, and the suggestion that this molecule is involved in re-setting the clock with light⁶. In support of this hypothesis, the *Drosophila* clock is also sensitive to blue light⁷. But the situation is somewhat more complicated. First, additional photopigments in the eyes seem to be critical for setting the clock. And second, clocks in most of the fly's organs are disrupted in *crybaby* mutants, indicating that cryptochromes may be involved in clock function⁶.

Do cryptochromes detect light in mammals? Initial data⁴ from *cry2* mutant mice suggest that the answer is yes (Table 1). Both *cry1* and *cry2* are expressed in the eye and in the suprachiasmatic nucleus (SCN) of the hypothalamus — the location of the 'master' clock in the mouse⁸. The induction by light of the mammalian *period 1* gene in the SCN is reduced in the *cry2* mutants, which show enhanced re-setting in response to pulses of light⁴. But some results do not fit well with this hypothesis. The wavelengths of light that are optimal for setting the circadian clock in mammals do not match the absorption wavelengths of cryptochromes. These wavelengths instead point to the involvement of another group of photopigments, the opsins⁹. Also, cryptochromes are present in the SCN, so if they were the circadian photopigments one would predict that mice without eyes could still set their circadian clocks in response to light. But it is well established that, in mice, the eyes are essential for light-dependent setting of the clock.

Indeed, the importance of the eye is confirmed by two reports in this week's *Science*^{10,11}, which show that mice lacking normal rods and cones — the typical light-responsive cells in the retina — still show normal re-setting of the clock in response to

Table 1 Evidence for the possible functions of cryptochromes in the circadian clock

Light-response function	Clock function
The normal light-dependent increase in transcription of the <i>period 1</i> gene is reduced in the <i>cry2</i> mutant mouse ⁴	Cryptochromes are expressed with a circadian rhythm in the retina and SCN, suggesting involvement in a central clock mechanism or associated pathway ⁸
The clock re-setting response to light is altered (increased) in the <i>cry2</i> mutant mouse ⁴	The effect of disrupting the <i>cry1</i> and <i>cry2</i> genes argues that cryptochromes are central in generating the rhythm ¹
Cryptochromes are found in retinal cells ⁹	Cryptochromes are found in the SCN ⁸
Argument by homology: role of cryptochromes in <i>Arabidopsis</i> and <i>Drosophila</i>	Opsin-based pigments seem to have a prominent part in the light response ⁹

light. Only when the authors surgically removed the eyes of these mice was this light-dependent entrainment blocked. These data indicate that a unique class of eye photoreceptors is essential for resetting the circadian clock.

The presence of cryptochromes in the SCN, the fact that they are expressed with a strong day–night rhythm, and the data from the mice lacking *cry1* and *cry2* are consistent with cryptochromes being central clock molecules. We now need to define the pathways through which the cryptochromes modulate known clock molecules, and so determine the exact molecular role they may play in clock function. □

David Whitmore and Paolo Sassone-Corsi are at the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS–INSERM–ULP, BP 163, 67404 Illkirch, Strasbourg, France.
e-mail: paolosc@igbmc.u-strasbg.fr

1. van der Horst, G. T. J. *et al. Nature* **398**, 627–630 (1999).
2. Todo, T. *et al. Science* **272**, 109–112 (1996).
3. Kobayashi, K. *et al. Nucleic Acids Res.* **26**, 5086–5092 (1998).
4. Thresher, R. J. *et al. Science* **282**, 1490–1494 (1998).
5. Somers, D. E., Devlin, P. F. & Kay, S. A. *Science* **282**, 1488–1490 (1998).
6. Stanewsky, R. *et al. Cell* **95**, 669–679 (1998).
7. Frank, K. D. & Zimmerman, W. F. *Science* **163**, 688–689 (1969).
8. Miyamoto, Y. & Sancar, A. *Proc. Natl Acad. Sci. USA* **95**, 6097–6102 (1998).
9. Foster, R. G. *Neuron* **20**, 829–832 (1998).
10. Freedman, M. S. *et al. Science* **284**, 502–504 (1999).
11. Lucas, R. J. *et al. Science* **284**, 505–507 (1999).

Galaxies

A long-distance record breaker

Richard Green

Direct observations of the earliest luminous objects are only now beginning to illuminate the study of the origin and formation of galaxies. Each step beyond the previous limit of discovery holds the promise that new techniques will reveal the precursors of the galaxies we see today. On page 586 of this issue, Chen, Lanzetta and Pascale¹ report the discovery of an object that probably has the largest redshift ever measured, $z = 6.68$. In the model of the expanding Universe, high redshifts — detected as shifts in spectral lines to longer wavelengths — of galaxies and quasars directly indicate their great distances from Earth. A redshift of 6.68 corresponds to a time when the Universe was only 5% of its present age. The previous record holder, reported last year², was found at a distance or ‘look-back’ time corresponding to when the Universe was 1% older.

Three factors have made the discovery of such high-redshift galaxies particularly difficult. First, such objects are not merely dimmed with distance by the cosmological equivalent to the inverse square law; galaxies are also extended, and surface brightness is dimmed by $(1 + z)^4$, so, only the brightest clumps in a system will be detectable above the sky and detector background. Second, there is no redshifted signal from the ultraviolet to the red part of the spectrum because the flux is completely absorbed by hydrogen in intergalactic clouds. Finally, the light produced and emitted in spectral lines by hydroxyl ions in the Earth’s own atmosphere can swamp faint signals emitted in the far red by the most distant objects.

Chen *et al.*¹ exploited the unique capability of the Space Telescope Imaging Spectrograph for discovering faint objects. The Hubble Space Telescope orbits above

the atmosphere and focuses images very sharply, allowing slitless spectroscopy in the far-red part of the spectrum. Potential high-redshift galaxies are distinguished from galaxies of lower redshift through differences in their spectral energy distributions.

The unique spectral signature of these proto-galaxies at high redshift is generated by a region of newly formed hot stars, which excite the remnant of the gas cloud out of which they were formed into emission, producing features such as the Lyman- α emission line in the ultraviolet spectrum of atomic hydrogen (Fig. 1). The Lyman- α line is a particularly strong line in the emission spectra of active galactic nuclei and quasars, which, if they are at high redshifts, appears shifted into the visible part of the spectrum. The combined ultraviolet radiation — the continuous spectrum from the hot stars and the emission lines from the gas — then propagates through an absorbing medium of more diffuse gas surrounding the region of star formation. The strong hydrogen absorption from this galactic gas, and from numerous other clouds along the line of sight, creates the final spectral signature of an absorption step and trough.

Has such a spectral signature been unambiguously identified in this case, to claim the current redshift record? Chen *et al.* have detected one emission line at near-infrared wavelengths of somewhat marginal significance, and the absorption step and trough is what you would expect from absorption in the galaxy and the intergalactic medium. The spectrum is consistent with the emission line being Lyman- α , but the redshift measurement is not yet of a quality to be widely and uncritically accepted as decisive. Traditionally, evidence of two spectral features of known wavelength

ratio or a spectral template match of relatively high fidelity is preferred. Some additional confirmation is required here for highest confidence, although the result seems very promising.

This candidate galaxy may be unusual, even among the few known very high-redshift objects, for two reasons. One is that the apparent ultraviolet luminosity is quite strong. On the basis of the known link between star formation rate and ultraviolet luminosity, the authors speculate that this result may indicate a trend to higher star-formation rates for individual galaxies, as we look back further in time. The other reason is that absorption from intergalactic hydrogen in the ultraviolet appears to be nearly 100% at this redshift of 6.68. At slightly lower redshifts, say 4.5, there is always some residual flux in the blue part of the emission spectrum. If the soup of intervening intergalactic clouds at $z \sim 6$ is largely composed of neutral gas, then the total hydrogen absorption would be opaque, as observed. To have some residual flux at a slightly lower redshift suggests that the clouds are partially or nearly totally ionized

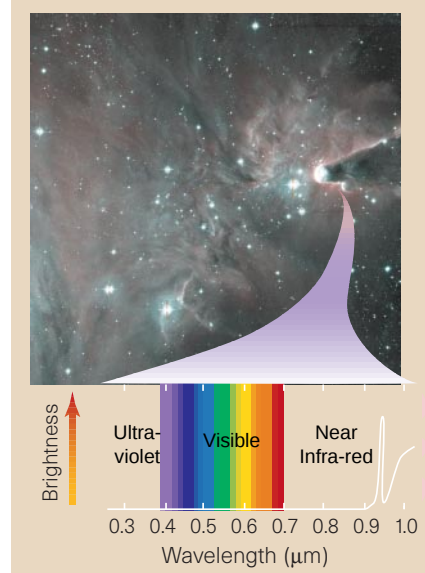


Figure 1 Expected spectral signature from a galaxy with a redshift of 6.68. Ultraviolet light from hot, young stars in a distant galaxy radiates out through a diffuse, absorbing cloud of gas across intergalactic space. The observed spectrum has a characteristic Lyman- α emission line, which is produced by hydrogen at $\sim 0.12 \mu\text{m}$ in the ultraviolet spectrum. The expansion of the Universe shifts the emission line to longer wavelengths. By the time it reaches us, the wavelength of light from an object at a redshift of z , has been stretched by a factor of $(1 + z)$. So Lyman- α emission from a galaxy at $z = 6.68$ appears to us as near-infrared radiation at $\sim 0.92 \mu\text{m}$ (see peak in spectrum). The dip in the spectrum just after the sharp peak is caused by strong absorption of the ultraviolet radiation by hydrogen in the galactic clouds and interstellar space.

KITT PEAK NATIONAL OBSERVATORY

over a very short interval of cosmic time. If total hydrogen absorption at $z \sim 6$ is confirmed with a higher signal-to-noise ratio observation, it would signal a sudden change in the properties of the intergalactic medium, perhaps by the 'switching on' of quasars and galactic star formation, both of which produce strong ionizing ultraviolet radiation.

The record for the highest redshift object in the known Universe is transient these days, and the differences in cosmic 'look-back' time for large jumps in redshift are very small. The significance of this discovery lies more in the promise of a systematic study of these most distant objects, of which only a handful are known today. There are critical questions waiting for answers. Most importantly, what are these objects and what is their evolutionary fate? Their spectra can be interpreted as revealing active star formation at the rate of 4–25 solar masses per year³. Such values are lower limits⁴ because of the obscuring effect of intervening dust. Brighter objects have low heavy-element content, implying that the starlight comes from one of the earliest generations of star formation⁵. The typical mass is still unknown, although it is likely to be sub-galactic. The suspicion is that these objects appear luminous because of strong star formation.

We observe these objects as they were in the past, but whether they will ultimately become typical spiral galaxies like our Milky Way, or the cores of massive elliptical galaxies is a mystery. The answer will help us choose between competing theories of galaxy formation. High-redshift objects ($z > 3$) may be regarded as the precursors to large, 'normal' galaxies because their approximate space density at $\sim 10\%$ of the current age of the Universe is consistent with the density of luminous present-day

galaxies⁵. Observations favouring such objects as the ancestors of elliptical cores are that many of the high-redshift galaxies are nearly spherical, and are unlikely to evolve into flattened rotating disks.

An argument for these ancient galaxies as precursors of spiral galaxies is that there are few close neighbours, which would be expected in the environment of a cluster where ellipticals are found⁶. Observed high-redshift objects are much more compact than present-day galaxies, so they might be sub-clumps that will ultimately merge. The objection to this picture is that you might expect to see small clusters of such objects, bound together and poised for coalescence. Typically, however, single objects have been detected. Perhaps such sub-clumps light up with newly formed stars at slightly different times, so that only one star-forming region per proto-galaxy is detectable at any very early epoch^{5,7}.

The advent of near-infrared spectroscopy with adaptive-optics — which compensate for the effects of the Earth's atmosphere — at ground-based telescopes, and stronger near-infrared capability in space, should put the study of the earliest luminous objects, now at the limit of our grasp, onto a sound footing. □

Richard Green is at the Kitt Peak National Observatory, Tucson, Arizona 85726-6732, USA.
e-mail: rgreen@noao.edu

1. Chen, H.-W., Lanzetta, K. M. & Pascarella, S. *Nature* **398**, 586–588 (1999).
2. Hu, E. M., Cowie, L. L. & McMahon, R. G. *Astrophys. J. Lett.* **502**, L99–L103 (1998).
3. Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M. & Adelberger, K. L. *Astrophys. J. Lett.* **462**, L17–L21 (1996).
4. Armus, L., Matthews, K., Neugebauer, G. & Soifer, B. T. *Astrophys. J. Lett.* **506**, L89–L92 (1998).
5. Lowenthal, J. D. et al. *Astrophys. J.* **481**, 673–688 (1997).
6. Giavalisco, M., Steidel, C. C. & Macchetto, F. D. *Astrophys. J.* **470**, 189–194 (1996).
7. Trager, S. C., Faber, S. M., Dressler, A. & Oemler, A. Jr *Astrophys. J.* **485**, 92–99 (1997).

El Niño

The child prodigy of 1997–98

Michael J. McPhaden

Just under a year ago, a sharp drop in equatorial Pacific sea-surface temperatures heralded the end of the 1997–98 El Niño. Called by some "the climate event of the century", it was by several measures the strongest on record.

Identifying why it was so strong challenges our understanding of the physical mechanisms responsible for El Niño. This is more than simply an academic question: the 1997–98 El Niño severely disrupted global weather patterns and Pacific marine ecosystems, and by one estimate caused \$33 billion in damage and cost 23,000 lives worldwide¹. There were warnings of an impending El Niño before it occurred. But

although many computer forecast models predicted 1997 would be warm in the tropical Pacific up to three seasons in advance, none predicted the rapid development or ultimate intensity of the event before its onset². Clearly we have much to learn from this experience.

El Niño, Spanish for 'the child' (and specifically the Christ child), is the name Peruvian fishermen gave to coastal sea-temperature warmings that first appeared around Christmas time. El Niño now more generally refers to a warming of the tropical Pacific basin that occurs roughly every three to seven years in association with a weakening of the trade winds. The flip side of El

Niño, La Niña, is characterized by stronger than normal trade winds and unusually cold sea-surface temperatures in the tropical Pacific. Both El Niño and La Niña are accompanied by swings in atmospheric pressure between the eastern and western Pacific known as the Southern Oscillation. These phenomena are collectively referred to as ENSO (El Niño/Southern Oscillation), which also includes periods of near-normal conditions³. At the moment, a strong La Niña is evident in the tropical Pacific, with several (but not all) forecast models predicting a return to normal by the end of 1999.

The general mechanisms underlying ENSO involve large-scale ocean-atmosphere interactions and equatorial ocean dynamics⁴. But each El Niño and La Niña is unique in the combination of its strength, duration and pattern of development. Irregularity in the ENSO cycle can be seen in both the instrumental record dating back to the middle of the last century, and in proxy data (such as lake sediments, coral growth rings, tree rings and ice cores) going back hundreds or even thousands of years⁵. So, in principle, it should not be surprising that an unusually strong El Niño occurs every so often.

Nonetheless, the 1997–98 El Niño was unprecedented⁶. It developed so rapidly that every month between June and December 1997 set a new monthly record high for sea-surface temperatures in the eastern equatorial Pacific⁷. December 1997 anomalies (that is, deviations from normal) were the highest ever recorded along the Equator in the eastern Pacific. Moreover, before 1997–98, the previous record-setting El Niño occurred in 1982–83. These two 'super El Niños' were separated by only 15 years, compared with a typical 30–40 year gap between such events earlier this century⁸.

Several factors may have contributed to the strength of the 1997–98 El Niño. One is chaos, which some theories invoke to account for the irregularity of the ENSO cycle. Nonlinear resonances involving ENSO and the seasonal cycle have received special attention, but other chaotic interactions may affect ENSO as well⁹. In 1997–98, events possibly conspired to produce an extraordinarily strong El Niño simply due to the underlying tendency towards chaos in the climate system. A related issue is that of weather 'noise'. Weather phenomena, inherently unpredictable more than about two weeks in advance, are a source of random forcing in the climate system. In the tropical Pacific, weather events occurring at the right time, and on time and space scales to which the ocean is sensitive, can markedly alter the evolution of the ENSO cycle¹⁰.

One notable source of weather in the

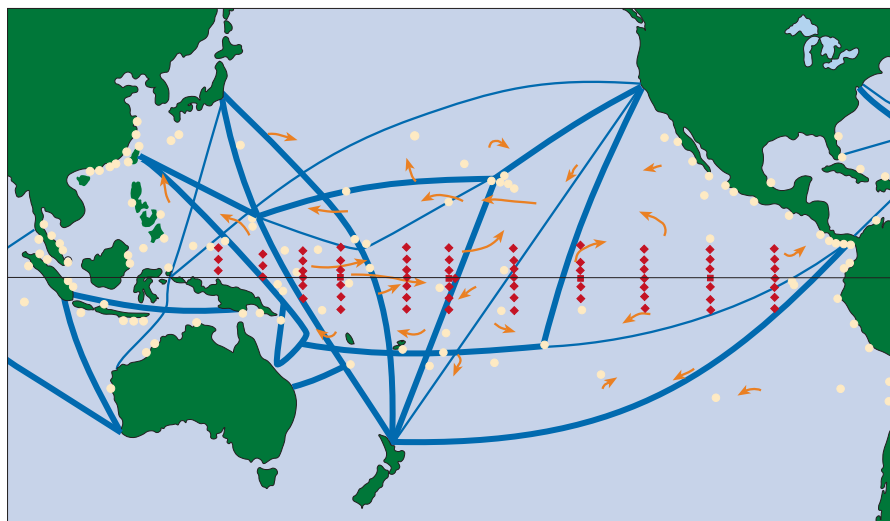


Figure 1 Taking the pulse of El Niño. The 1997–98 event was followed through the El Niño/Southern Oscillation (ENSO) Observing System, which was set up to monitor and predict ENSO variations. The Earth-based components, shown here, largely relay data in real time via satellites. The main components are a volunteer ship programme (blue tracks); an island and coastal tide-gauge network (cream); and systems of drifting (orange arrows) and moored buoys (red). Complementing this network are satellites that provide data from space with near-global coverage. They include the US/French TOPEX/Poseidon mission; the European Space Agency Earth Remote Sensing satellites; US Department of Defense satellites; and NOAA's polar-orbiting weather satellites. Taken together, this ensemble of instrumentation delivers data on surface and subsurface temperature, wind speed and direction, sea level, and current velocity. The ENSO Observing System was completed in 1994 at the end of the ten-year international Tropical Ocean Global Atmosphere programme. It is now being continued in support of operational climate forecasting as well as research on ENSO dynamics.

tropics is the Madden–Julian Oscillation (MJO), a wave-like disturbance in the atmosphere with a period of 30–60 days that originates over the Indian Ocean¹¹. It could have been that the ocean got a healthy kick from the MJO at just the right time to send it on a course towards record high temperatures⁷. The tropical Pacific was preconditioned to the onset of an El Niño by the build-up of excess heat in the western equatorial Pacific due to stronger than normal trade winds in 1995–96. However, beginning in late 1996, the MJO was particularly energetic, and several cycles of the wave amplified through nonlinear ocean–atmosphere interactions as they passed over the western Pacific. This set in motion a series of positive feedbacks between the ocean and the atmosphere which reinforced initial MJO-induced warming.

Another possibility is that the ENSO cycle may be interacting with the Pacific Decadal Oscillation (PDO) — which, as the name implies, is a naturally occurring oscillation of the coupled ocean–atmosphere system in the Pacific basin with a period of several decades¹². In association with the PDO, sea-surface temperatures have generally been higher in the tropical Pacific from the mid-1970s. Since then, there have been more El Niños than La Niñas, the early 1990s was a period of extended warmth in the tropical Pacific, and two super El Niños have occurred. The PDO may be one of the

reasons for the observed decadal modulation of the ENSO cycle, because it affects the background conditions on which ENSO events develop. From that perspective, the strength of the 1997–98 El Niño may be but one manifestation of a linkage between interannual and decadal climate variations in the Pacific.

Global warming trends are yet another possible influence on the ENSO cycle. The warmest years on record were, in order, 1998 and 1997. The 1997–98 El Niño contributed in part to these record highs, because global mean temperatures generally rise a few tenths of a degree Celsius following the peak El Niño warming as the tropical Pacific loses heat to the overlying atmosphere^{13,14}. Underlying these extreme temperatures, however, is a century-long warming trend that may well be due to anthropogenic greenhouse-gas warming¹⁵.

Some computer models suggest that global warming may be slowly heating up the eastern equatorial Pacific Ocean, as observed over the past 25 years¹⁶. Others propose that ENSO events may be stronger or more frequent in a warmer climate¹⁷. The superposition of ENSO variations on increased warming due to CO₂ and a warm phase of the PDO could produce temperature fluctuations like those seen in the equatorial Pacific since the mid-1970s, including the extreme temperatures associated with the 1997–98 El Niño¹⁸. But computer models used in global change studies are limited



100 YEARS AGO

What constitutes the natural prey of the lion in his wild state is, I believe, a disputed point. The majority of people, probably, are of opinion that he is extremely fastidious in his tastes; others, again, assert that he will eat almost anything. Certainly, it is only reasonable to suppose that a lion sufficiently under the impulse of hunger will eat “almost anything”! Years ago I was present on more than one occasion when animated discussions on this point took place between two notable African ecclesiastics — both since dead — Bishop Smythies and Archdeacon Maples (he was then), both of whom had travelled a good deal in Africa — Maples more especially — and had seen something of the habits of lions. Bishop Smythies defended the former theory; Archdeacon Maples — a most talented and entertaining man — the latter, saying he had known instances of lions killing porcupines, and adding that he believed the porcupine to be specially endowed with the power to propel his quills into his assailant when so attacked. At this juncture, Bishop Smythies generally lost patience and declined to continue the argument. Had Bishop Smythies lived, it would have interested him ... to know that in March last, at the Salt Stream, two days' march N. W. of Kibwezi, I shot a fine old lion in whose left fore-paw were deeply buried the tips of three porcupine quills.

From *Nature* 13 April 1899.

50 YEARS AGO

It is announced that Prof. J. W. McBain, who has recently retired from the chair of physical chemistry at Leland Stanford University, California, has been appointed the first director of the National Chemical laboratory of India. It is difficult to think of a better choice. While his many English friends have not seen so much of him since he left the University of Bristol, yet they have been able to follow his steady development of the concept of the micellar structure of colloidal electrolytes, a chapter in physical chemistry which is particularly his own. ... We must recollect that we are also indebted to him for coining the word ‘sorption’, thus directing attention to the complexity of the interactions between a gas and a solid.

From *Nature* 16 April 1949.

in their ability to represent many key processes involving physical, chemical and biological interactions of the coupled ocean-atmosphere-land system^{15,19}. So although we might expect global warming to affect the ENSO cycle, at the moment no firm conclusions can be drawn.

In short, the strength of the 1997–98 El Niño may be traceable to interactions of the ENSO cycle with some combination of global warming trends, the Pacific Decadal Oscillation, the seasonal cycle and the Madden-Julian Oscillation. Complicating the picture is a general background of weather noise and chaotic interactions in the climate system. Yet unconsidered processes may also be involved.

This complexity may be daunting, but there is a bright side. The 1997–98 El Niño was not only among the strongest on record, it was also the most comprehensively observed²⁰ (Fig. 1). The 'climate event of the century' may be over, but measurements from the recently completed ENSO observing system will stimulate research into the causes and consequences of El Niño for years to come. The hope is that we will ultimately be able to unravel the various threads that were woven together to produce the 1997–98 El Niño, and to translate that knowledge of interactions spanning

intraseasonal to centennial timescales into better models for climate forecasting. □

Michael J. McPhaden is in the NOAA/Pacific Marine Environmental Laboratory, 7600 Sand Point Way NE, Seattle, Washington 98115, USA.
e-mail: mcphaden@pmel.noaa.gov

1. Kerr, R. *Science* **283**, 1108–1109 (1999).
2. Barnston, A. G., Glantz, M. H. & He, Y. *Bull. Am. Meteorol. Soc.* **80**, 217–243 (1999).
3. Trenberth, K. E. *Bull. Am. Meteorol. Soc.* **78**, 2771–2777 (1997).
4. Neelin, J. D. *et al.* *J. Geophys. Res.* **103**, 14261–14290 (1998).
5. Rodbell, D. T. *et al.* *Science* **283**, 516–520 (1999).
6. Webster, P. J. & Palmer, T. N. *Nature* **390**, 562–564 (1997).
7. McPhaden, M. J. *Science* **283**, 950–954 (1999).
8. Caron, J. M. & O'Brien, J. J. *Mon. Weath. Rev.* **126**, 2809–2821 (1998).
9. Tziperman, E., Stone, L., Jarosh, H. & Cane, M. A. *Science* **264**, 72–74 (1994).
10. Chen, Y.-Q., Battisti, D. S., Palmer, T. N., Barsugli, J. & Sarachik, E. S. *Mon. Weath. Rev.* **125**, 831–845 (1997).
11. Madden, R. A. & Julian, P. R. *J. Atmos. Sci.* **29**, 1109–1123 (1972).
12. Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M. & Francis, R. C. *Bull. Am. Meteorol. Soc.* **78**, 1069–1079 (1997).
13. Christy, J. R. & McNider, R. T. *Nature* **367**, 325 (1994).
14. Schneider, E. K., Lindzen, R. S. & Kirtman, B. P. *J. Atmos. Sci.* **54**, 1349–1358 (1997).
15. Intergovernmental Panel on Climate Change *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) (Cambridge Univ. Press, 1996).
16. Meehl, G. A. & Washington, W. M. *Nature* **382**, 56–60 (1996).
17. Timmerman, A. *et al.* *Nature* (in the press).
18. Lau, K.-M. & Weng, H. J. *J. Clim.* (in the press).
19. Feely, R. A., Wanninkhof, R., Takahashi, T. & Tans, P. *Nature* **398**, 597–601 (1999).
20. McPhaden, M. J. *et al.* *J. Geophys. Res.* **103**, 14169–14240 (1998).

Malaria

A new escape and evasion tactic

John W. Barnwell

The mosquito-borne parasites (*Plasmodium*) that cause malaria have adapted a variety of devious molecular strategies to enable them to survive and keep propagating in their immunologically aggressive hosts. Now, on page 618 of this issue, Preiser *et al.*¹ reveal another clever new weapon that has been developed by *Plasmodium* in order to escape destruction in the molecular arms race between host and parasite.

Malarial disease results from the cycles of multiplication of asexual *Plasmodium* parasites in the blood of the host. This cyclical propagation is initiated by an invasive form of the parasite, the merozoite, when it enters the red blood cells (erythrocytes). Once inside, merozoites undergo transformation into ring-shaped trophozoites, which mature into large trophozoites before undergoing schizogony, a form of haploid nuclear division. This in turn generates between six and 32 new merozoites, which burst out of their erythrocyte casement into the bloodstream.

Merozoites do not live for long outside the cell and so must quickly find and enter another erythrocyte. This recognition and invasion of host cells is executed through a 'sensor' of polypeptide components, which

are largely located in a complex of organelles (termed rhoptries and micronemes) at the distinctive apical pole of the merozoite².

Because the merozoite components that mount the invasion represent antigens that could potentially be used for developing a vaccine against malaria, it is important to understand the mechanisms that have evolved in malaria parasites to avoid a host blockade against invasion. In theory, malaria parasites should be vulnerable to neutralizing antibody responses mounted by the host against any of the critical merozoite proteins that enable them to invade the host cell. Also, any alteration in a complementary receptor on the target host cell could disrupt the interaction and prevent the merozoite from entering it.

Like many other pathogens, malarial parasites have evolved mechanisms to overcome such tactics. One used by *P. falciparum*, which causes a severe malaria in humans, and by *P. knowlesi*, a monkey parasite, is to express up to three alternative 'sticky' adhesin molecules in merozoites. These adhesins all help the merozoite to invade its target host cell, but they each engage different receptors on the erythrocyte^{3,4}.

Preiser *et al.*¹ have found evidence for an

unusual and fascinating mode of clonal phenotypic variation used by another *Plasmodium* species, *P. yoelii* to avoid the barricades of its host, a small African rodent (*Thamnomys rutilans*). Clonal antigenic variation of parasite-derived proteins presented at the surface of occupied erythrocytes is a known phenomenon in malaria^{5–7}. The unique form of molecular variation described by Preiser *et al.*¹ takes place not in the intracellular vegetative trophozoite, however, but during the development of the invasive merozoite parasite form (see their Fig. 3 on page 620).

Preiser *et al.* have investigated the pattern of transcription for a large multigene family that expresses antigenically variable proteins, known collectively as p235, which localize in the rhoptry organelles at the apical pole of *P. yoelii* merozoites. There are at least 11, and probably up to 50, genes encoding discrete and variable members in the family⁸. Active immunization experiments, passive transfer of antibody and binding of at least one p235 protein to mature erythrocytes have indicated that these proteins play a role in selecting erythrocytes for merozoite invasion⁹.

Based upon their analysis of p235 transcripts in individual parasites at different stages of development, Preiser *et al.*¹ propose that a single p235 gene is transcribed in each nucleus that is replicated during schizogony. This by itself is not unusual. But in this case, what is unusual is that the particular p235 gene transcribed can be different for each nucleus present in a single developing schizont. The authors surmise from this that each of the individual merozoite progeny derived from a parent schizont could express a different p235 protein.

Why would the parasite use this method of phenotypic variation? It is presumed (but not yet confirmed) that the specificity of p235 proteins for their target receptors is variable, so a single mature schizont could give rise to a mosaic of merozoites, each possessing unique receptor specificity, or at least antigenicity. Accordingly, this could afford the parasite some form of insurance against the possibility of all the progeny (merozoites) of a schizont being destroyed by immune responses specific for some p235 proteins but not others. Survival advantage could also be accrued if this pattern of receptor expression maximized the benefits of differences in the specificity of host-cell recognition. Expansion into new host niches would be facilitated, and parasitaemias could be maintained in the face of a changing haematological landscape where erythrocytes of a certain maturity were either increasing or decreasing in number.

Although it is not the natural host of *P. yoelii*, the white laboratory mouse can be infected by this parasite, making it a popular model for investigating malaria. In this context, there are two relevant questions. First, is

there a similar gene family present in any of the four species of *Plasmodium* that cause malaria in humans? Second, does a similar type of transcriptionally mediated phenotypic variation occur in human malaria parasites?

Plasmodium vivax, a major cause of malaria morbidity in humans, expresses a 'sensor' complex containing two adhesive proteins (RBP-1 and RBP-2) that probably assists its merozoites in invading young red blood cells (reticulocytes)¹⁰. The p235 rhoptry proteins show a low but significant level of identity and structural similarity with RBP-2 (refs 11,12). However, RBP-2 in *P. vivax* is a single-copy gene, not a member of a gene family¹⁰. The other primary cause of human malaria, *P. falciparum*, has two genes which express proteins that are distantly related to the *P. yoelii* p235 family and *P. vivax* RBP-2 (our unpublished data). Whether the other two species that cause malaria, *P. malariae* and *P. ovale*, also express merozoite proteins similar to p235 is not known.

This suggests that neither *P. vivax* nor *P. falciparum* has developed an extensive gene family from their p235-homologue genes in the way that *P. yoelii* did for the p235 rhoptry proteins, and so these parasites must have followed a different evolutionary route. This is not really surprising as the genus *Plasmodium* is not monophyletic¹³ — *P. yoelii*, *P. vivax* and *P. falciparum* are not closely related, having diverged from each other many tens of millions of years ago. In the process, these parasites may have adapted individual mechanisms of evasion: *P. yoelii* by developing a repertoire of p235 genes and choosing different ones for expression in daughter merozoites, and *P. falciparum* by targeting alternative host-cell receptors to function at different steps in the invasion process.

No comparable mechanism of phenotypic variation has so far been discovered in *P. vivax* or *P. falciparum* merozoites, although these species readily establish long-term chronic infections. But it is conceivable that the type of variation identified by Preiser *et al.*¹ in the *P. yoelii* p235 family might operate in other protein families in the different species of *Plasmodium* that infect humans. If it does, the race to outwit the parasite will take a new turn. □

John W. Barnwell is at the Centers for Disease Control and Prevention, Division of Parasitic Diseases, 4770 Buford Highway NE, Atlanta, Georgia 30341, USA.
e-mail: w2b3@cdc.gov

1. Preiser, P. R., Jarra, W., Capoid, T. & Snounou, G. *Nature* **398**, 618–622 (1999).
2. Bannister, L. H. & Dlugewski, A. R. *Blood Cells* **16**, 257–292 (1990).
3. Dolan, S. A. *et al. Mol. Biochem. Parasitol.* **64**, 55–63 (1994).
4. Adams, J. H. *et al. Proc. Natl Acad. Sci. USA* **89**, 7085–7089 (1992).
5. Baruch, D. I. *et al. Cell* **82**, 77–87 (1995).
6. Su, X. Z. *et al. Cell* **82**, 89–100 (1995).
7. Al-Khedery, B., Barnwell, J. W. & Galinski, M. R. *Mol. Cell* **3**, 131–141 (1999).

8. Borre, M. B. *et al. Mol. Biochem. Parasitol.* **70**, 149–155 (1995).
9. Ogun, S. A. & Holder, A. A. *Mol. Biochem. Parasitol.* **76**, 329–332 (1996).
10. Galinski, M. R., Corredor-Medina, C., Ingravallo, P. & Barnwell, J. W. *Cell* **69**, 1213–1226 (1992).

11. Keen, J. K. *et al. Mol. Biochem. Parasitol.* **65**, 171–177 (1994).
12. Galinski, M. R. & Barnwell, J. W. *Parasitol. Today* **12**, 20–29 (1996).
13. Escalante, A. A. & Ayala, F. J. *Proc. Natl Acad. Sci. USA* **91**, 11373–11377 (1994).

Earth science

Screws tighten on the core

Michael J. Walter

A popular model for the formation of Earth's iron-rich core is that metal equilibrates with molten silicate in a deep magma ocean at an average pressure of about 28 GPa (that is, at a depth of 800 km or so). This model can apparently account for the magnitude of depletion of many siderophile (metal-loving) elements in Earth's upper mantle, most notably that of Co and Ni^{1,2}.

On page 604 of this issue³, Tschauner *et al.* present new experimental data for partitioning of Co and Ni between metal and the most abundant mineral in the lower mantle, Mg-perovskite. Their data indicate that metal–silicate equilibration would cause an enrichment in these elements in the lower mantle relative to the upper mantle, and this may require some modification of the core-formation model that invokes a magma ocean.

Siderophile elements such as Co, Ni, W, Re and Au partitioned strongly into the Fe-

rich metal that segregated to form Earth's core. Consequently, these elements are depleted by up to orders of magnitude in Earth's silicate upper mantle relative to their abundance in the early Solar System (as inferred from certain chondritic meteorites). The amount by which each siderophile element is depleted strongly depends on intensive variables such as temperature, pressure and overall composition. To provide the information necessary for modelling the observed depletions, experimentalists measure metal/silicate partition coefficients, $D^{\text{met/sil}}$, over a wide range of conditions.

Two elements that are especially important are Co and Ni, because the Co/Ni ratio in Earth's upper mantle is nearly the same as in chondrites, indicating that they were not fractionated from one another during core formation. This means that, on average, the ratio of their partition coefficients, $D^{\text{Ni}}/D^{\text{Co}}$, was nearly unity, placing a power-

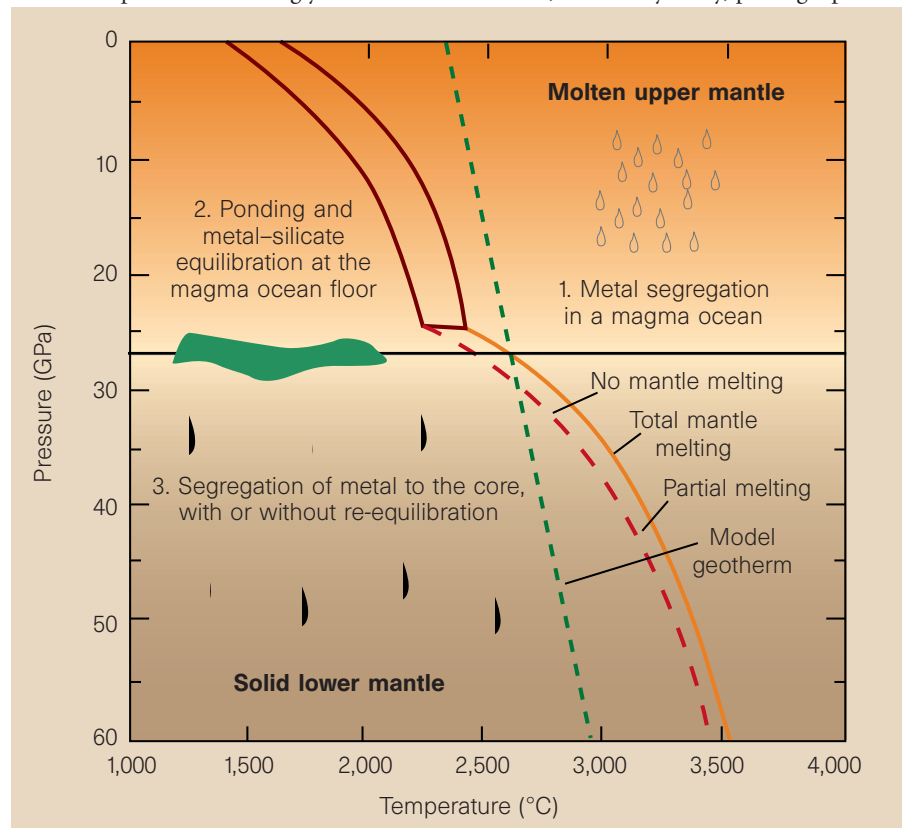


Figure 1 Model for the segregation of Earth's metal core from a deep magma ocean. Metal segregates from the molten upper mantle, and ponds and equilibrates at the magma ocean floor. The metal then migrates through the solid lower mantle, where some re-equilibration may take place.

ful constraint on core-formation models. Experimental studies have revealed that the combined effects of pressure and temperature conspire to produce conditions that can nearly satisfy the absolute and relative abundances of Co and Ni in the upper mantle by equilibration between molten silicate and molten metal at about 28 GPa and 2,300 K. These conditions imply metal segregation in a deep magma ocean^{1,2}.

A pressure of 28 GPa corresponds closely to the depth at which a peridotite mantle transforms from an upper mantle dominated by olivine and its polymorphs, to a lower mantle dominated by silicate perovskite. This mineral transition affects the melting behaviour of the mantle such that the melting curve becomes much steeper at pressures above the transition⁴. The high melting temperature of a perovskite-rich lower mantle means that the upper mantle may have been molten while the lower mantle was solid at the time of core formation; that is, metal and silicate could have equilibrated at the floor of a deep magma ocean as depicted in Fig. 1.

Modern theory for the origin of Earth has it that most of its mass accreted as a consequence of catastrophic impacts with large planetesimals⁵. Indeed, a popular model for the origin of the Moon is that it formed when a Mars-sized body hit the proto-Earth. This purported 'giant' impact would have occurred late in Earth's accretionary stage, and both impactor and Earth probably already had dense metal proto-cores. Numerical models indicate that much of the impactor core may have ploughed its way through the mantle and merged with Earth's core. Concurrently, the kinetic energy buried deep within the Earth would have been sufficient to melt a large portion of the silicate mantle^{6,7}.

Although the complex dynamics of this event are still being explored theoretically, it could have been that metal from the pre-impact cores of the proto-planets was excavated, re-mixed and re-equilibrated with the molten mantle. The Co and Ni contents of the upper mantle indicate that metal-silicate equilibration occurred at the base of the molten upper mantle; so, to form a core, the metal would have to migrate through the solid lower mantle (Fig. 1). During this final segregation stage, metal may have re-equilibrated to some degree with the major minerals of the lower mantle, silicate perovskite (about 85%) and magnesiowüstite (about 15%). An unknown factor is the effect this would have had on mantle abundances of Co and Ni.

In their provocative study, Tschauner *et al.*³ used a powerful experimental and analytical combination — the laser-heated diamond anvil cell and secondary ion microprobe — to measure the partitioning behaviour of Co and Ni between Mg-per-

ovskite and Fe-rich metal at lower-mantle pressures (up to 80 GPa, corresponding to a depth of around 1,900 km). These extremely difficult experiments provide the first measurements of metal-silicate partitioning over a range of lower mantle conditions, and the results indicate that the partition coefficients for Co and Ni both approach unity at very high pressure, and that $D^{\text{Ni}}/D^{\text{Co}}$ is about one throughout the lower mantle. The implication of these data is that any re-equilibration between metal and silicate during migration to the core would increase the absolute abundance of Co and Ni in the lower mantle relative to the upper mantle, without fractionating the Co/Ni ratio.

For example, assuming a core that constitutes 30% of Earth's mass, metal-silicate partition coefficients of 42 and 46 for Co and Ni, respectively, are required to account for their upper mantle depletions. So partition coefficients of less than these values, and which approach unity in the lower mantle, would cause the silicate to be enriched in Co and Ni relative to observed depletions. Because the magma ocean model already accounts for the Co and Ni abundances, this creates something of a budgetary crisis.

The dilemma may be circumvented in three ways: by assuming that there was no re-equilibration in the lower mantle; by attempting to match the mantle Co and Ni depletions by developing a new model combining both upper mantle and lower mantle equilibration; and by assuming that a Co- and Ni-enriched lower mantle has remained chemically isolated for the past 4.5 billion years.

The study by Tschauner *et al.*³ is the first of its kind at ultra-high pressures, and the implications of their work demonstrate the need for continued exploration into metal-silicate partitioning under lower-mantle conditions. For example, another unknown in core-formation models is the effect of metal-magnesiowüstite partitioning, as this mineral probably comprises about 15% of the lower mantle. Such studies can yield insights not only into the ancient process of core formation, but also the possible reaction of the outer core with lower-mantle minerals throughout geological time. □

Michael J. Walter is at the Institute for Study of Earth's Interior, Okayama University, Misasa, Tottori-Ken, 682-01, Japan.
e-mail: walter@misasa.okayama-u.ac.jp

- Li, J. & Agee, C. *Nature* **381**, 686–689 (1996).
- Righter, K., Drake, M. & Yaxley, G. *Phys. Earth Planet. Int.* **100**, 115–134 (1997).
- Tschauner, O. *et al.* *Nature* **398**, 604–607 (1999).
- Zerr, A., Diegeler, A. & Boehler, R. *Science* **281**, 243–246 (1998).
- Wetherill, G. W. *Science* **228**, 877–879 (1985).
- Benz, W. & Cameron, A. G. W. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 61–68 (Oxford Univ. Press, New York, 1990).
- Melosh, H. J. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 69–83 (Oxford Univ. Press, New York, 1990).

Daedalus

Slow-flowing metal

The whole of metallurgy is determined by the fact that molten metals have a very low viscosity. Casting them is easy and simple. Some metals can be forged as solids, by raising them to a temperature at which they can readily be deformed. But molten metals are not easily blown or drawn like glass, or spread like tarmac, or vacuum-formed like plastics. Such processes would need a metallic melt that was viscous and tacky. So Daedalus is inventing one.

What is needed, he says, is some sort of long-chain polymer that will dissolve in the molten metal without decomposing. Even in very low concentrations, it would thicken the melt dramatically, as linear polymers do in solution, or fine fibres in suspension. At first sight the chemistry looks discouraging. Organic polymers decompose at the temperature of most molten metals, and silicones would be immiscible with them. Some sort of long-chain silicone molecule with metallic side-chains might be feasible. But the obvious answer is that modern solution in search of a problem, the carbon nanotube.

Many alloys, especially those of iron, contain carbon inclusions, so problems of compatibility should not arise. Carbon nanotubes have a high (indeed, unknown) melting point. They have been filled with molten metal, so are clearly wetted by it; they can be made in micrometre sizes like chain-polymer molecules, and they are immensely strong. They are certainly expensive, but are rapidly getting cheaper. In any case only a small concentration of them should transform a runny molten metal into a viscous liquid. The resulting product should transform metallurgy.

The most immediate application would be to those shell structures — cans, boxes, containers, car-bodies and so on — which are now pressed or fabricated from sheet. An alloy which could be blow-moulded like glass would make it possible to form such products in one simple, effortless operation. Similarly, wire could be drawn in the finest sizes by simple melt-pulling, like glass fibre. Thin-wall metal tube and sheet could be blown continuously, like plastics layflat tubing. Light, strong metal foam or insulating and resilient metal wool could be made by blowing gas into the sticky melt at a well-judged velocity. Better still, all the resulting products could easily be repaired or modified with a simple blow-torch. Heated to tackiness, they could be annealed, augmented, deformed, or teased into new shapes by the skills of the traditional glassblower.

David Jones

Obituary

Thomas A. McMahon (1943–99)

Biomechanist who discovered fundamental principles of biological structure and function

Thomas A. McMahon, who died suddenly on 14 February, was a pioneer in the field of biomechanics, a discipline that integrates engineering and biology. During his 30-year career at Harvard University, McMahon sought the fundamental mechanical principles that determine how organisms are built and work. Today, most biomechanists specialize in one or maybe two sub-disciplines, such as the mechanics of fluids, plants, bones, muscles or locomotion. McMahon contributed to each of these areas — not as a dabbler, but as a penetrating thinker who identified the key mechanical parameters that shape biological structure and function.

McMahon liked to integrate his hobbies and science. He enjoyed rowing his single shell on rivers and lakes. After observing that larger shells with more rowers could glide past him at high speed, he developed a simple mechanical model to explain why. McMahon's rowing project catapulted him into the world of biological scaling, a transition that led him to develop the theory of 'elastic similarity'. The guiding principle of this theory is that organisms of different sizes are built so that their structures experience similar elastic deformations in their everyday lives. McMahon measured the dimensions of animal bones and trees to see how plausible his theory was. He spent a sabbatical in Vermont, dismantling and measuring the dimensions of an entire tree, and discovered that longer tree limbs are relatively stouter than shorter limbs. As a result, all limbs droop to a similar angle under their own weight — as predicted by the elastic similarity theory.

In 1969, McMahon met the late C. Richard Taylor, a fellow Harvard professor and comparative physiologist, who profoundly influenced McMahon's career by introducing him to the mysteries of animal locomotion. McMahon's interactions with Taylor were extraordinarily fruitful because their scientific styles differed dramatically. When faced with a difficult question about how animals work, McMahon pulled out a piece of paper and derived a mathematical model to answer the question. In contrast, Taylor designed an elegant experiment. Although McMahon and Taylor did not publish many papers together, they shaped the scientific styles of their students, many



of whom learned the power of bringing together models and experimentation. McMahon's view of how mathematical modelling can be used to understand locomotion is demonstrated in his influential book *Muscles, Reflexes and Locomotion* (Princeton Univ. Press, 1984).

After turning his attention to locomotion, McMahon revolutionized our understanding of walking and running by emphasizing how the mechanical properties of the body influence locomotor movements. McMahon and Simon Mochan developed a 'ballistic walking' model, which demonstrated that the swing limb of a walking human behaves like a passive compound pendulum, and that it is coupled to the stance limb, which behaves like an inverted pendulum. They concluded that the coupling of these pendula is critical in determining the walking motion. The plausibility of this idea was later proven when Tad McGeer invented a passive robot that walked by using these pendular mechanisms.

McMahon continued to show that the body's mechanical properties are crucial for shaping locomotion, when he worked out how the spring-like stiffness of the legs determines running dynamics. This discovery led to the invention of the 'tuned track'. McMahon and Peter R. Greene realized that, by matching the stiffness of the track to that of the human leg, they could increase running speed. Indeed, after their tuned track was installed at Harvard, performances in the quarter mile and longer races improved by about 3%. Later, work by McMahon and others revealed that this spring-like behaviour of limbs might be an organizing principle of the neuromuscular system when animals run, hop or trot.

Locomotion biomechanics and fluid mechanics are usually separate sub-disciplines of biomechanics, but McMahon

and James W. Glasheen combined them to explain how basilisk lizards — also called 'Jesus Christ' lizards — can run on the surface of water. These lizards use three tricks to do this. First, the lizard's fringe-toed foot slaps the water surface, and then it rapidly strokes downward creating an air cavity. Both the slap and stroke contribute upward impulse, preventing the lizard from sinking. Finally, the lizard pulls its foot upward before the air cavity collapses, so minimizing the downward drag force on the foot.

McMahon enjoyed inventing practical devices that were inspired by his biomechanical knowledge. Recently, with Wilson C. Hayes and Stephen N. Robinovitch, he designed an ingenious device to reduce the risk of hip fractures when elderly people fall. They invented a U-shaped hip pad that surrounds the head of the femur. Composed of a material that is normally liquid, this pad conforms to the body during the slow movements that occur during everyday activities. But during the impact of a fall, rapid deformation of the material stiffens it into a solid, which shunts force away from the head of the femur and into the surrounding muscles and fat.

McMahon told his friends that he had "the best job in the world", because his duties included discovering how nature works and interacting with bright young students. His reputation as a thoughtful and caring mentor led Harvard to install a bench outside his office door to accommodate the perpetual queue of students seeking guidance. McMahon was modest, unconventional and playful. His golden retriever, Tess, accompanied him everywhere, even to traditional social affairs on the Harvard campus. He often invited students to join him for "fetch breaks" — respites from academic work, to play with his canine companion on the lawn outside his office. McMahon was also an acclaimed novelist. He wrote three novels, all of which featured scientists, engineers or inventors as the main characters.

McMahon wrote about his view of the interplay between scientific progress and nature's mysteries: "Theories rarely threaten the mystery of natural phenomena. The best theories harmonize with that mystery and make it ever more fascinating". Indeed, McMahon's melding of engineering and biology to discover how organisms work did just that.

Claire T. Farley

Claire T. Farley is in the Department of Integrative Biology, University of California, 3060 Valley Life Sciences Building, Berkeley, California 94720-3140, USA.

e-mail: cfarley@socrates.berkeley.edu

Drug addiction: bad habits add up

Trevor W. Robbins and Barry J. Everitt

We know how many drugs of abuse — cocaine, heroin and nicotine — work, but less about how they lead to addiction. Studies of the brain-learning systems concerned are addressing the causes of addiction, with the intent of developing better treatments.

Drug addiction — which is increasingly seen as a neuropsychiatric disorder — places an enormous burden on society through its repercussions on crime rate and healthcare. The economic costs of addiction have been estimated at 80 billion dollars in the United States alone, and many Western countries have invested heavily in research towards understanding, treating and preventing addiction. Recently, through genetic and cell-biological approaches, many of the molecular targets for drugs of abuse have been identified and cloned. But the value of these powerful reductionist approaches depends, in turn, on an integrative framework of systems and cognitive neuroscience. Such a framework allows us to formulate new hypotheses that take into account the complex factors influencing addiction and its treatment.

The dopamine hypothesis

By the early 1990s, converging evidence suggested that many (if not all) drugs of abuse act through mechanisms involving the brain neurotransmitter dopamine and the neural systems that it regulates¹. Although these drugs — including 'stimulants' such as amphetamine and cocaine, opiates such as heroin, and even 'legal' drugs such as alcohol and nicotine (Box 1, overleaf) — can influence several different chemical neurotransmitter systems in the brain, many of these 'primary' responses lead to secondary effects involving dopamine. For example, morphine and heroin bind first to an opiate receptor, which then increases the activity of the so-called 'mesolimbic' dopamine neurons in the midbrain. These neurons send their projections to interconnected forebrain structures such as the prefrontal cortex and striatum (Fig. 1). A region at the base of the striatum, the nucleus accumbens, is the key zone that mediates the rewarding effects of drugs such as amphetamine and cocaine, which act directly by increasing the levels of dopamine at this site.

Evidence for this hypothesis came from several sources¹. Rats will self-administer tiny injections of amphetamine to the nucleus accumbens (by pressing a lever to activate a microsyringe connected to a stainless-steel tube, or cannula, implanted there)². The rats

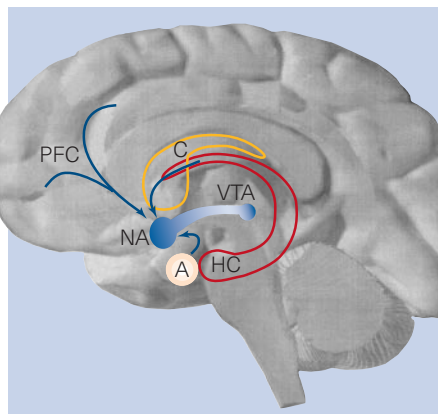


Figure 1 Some of the brain structures affected by drugs of abuse. The mesolimbic dopamine system originates in the ventral tegmental area (VTA) of the midbrain, and projects to the nucleus accumbens (NA). The amygdala (A), hippocampus (HC) and medial prefrontal cortex (PFC) send excitatory projections to the nucleus accumbens. C, caudate nucleus (striatum).

give themselves more of the amphetamine when their dopamine receptors are partly blocked pharmacologically, suggesting a drive to self-regulate the level of dopamine activity (possibly a form of homeostasis; see Box 2 on page 570 for terminology). This tendency to repeat behavioural acts that lead to drug effects we assume to be rewarding is an example of 'positive reinforcement' — a necessary feature of certain forms of memory or learning. If dopamine is massively depleted using a neurotoxin called 6-hydroxydopamine then, despite the recuperative capacity of the dopamine systems, the rats no longer self-administer amphetamine and cocaine. This is presumably because, in the absence of dopamine, these drugs lose their reinforcing properties¹.

The rewarding effects of drugs other than stimulants may also depend on the mesolimbic dopamine system. For example, rats will also self-administer morphine to the area of the midbrain from which the dopamine neurons project³. Using a microdialysis technique, which allows extracellular dopamine concentrations to be sampled directly from living brain tissue, withdrawal from several drugs — including alcohol and nicotine, as well as opiates and stimulants — has been

shown to be associated with reduced levels of dopamine in the nucleus accumbens⁴. From these observations comes the 'modal hypothesis', which states that the reinforcing effects of all drugs of abuse partly depend on the mesolimbic dopamine system. These effects could, perhaps, stem from the more general role of this system in mediating the motivating properties of natural stimuli such as food or sex. In some sense, then, drugs 'short-circuit' or 'usurp' normal behavioural and motivational processes mediated by this region of the brain¹. But although there is compelling evidence that amphetamine-like stimulants interact with this circuitry, the 'strong' form of the dopamine hypothesis — which embraces all other drugs of abuse — is certainly not universally accepted. For example, opiates also seem to have reinforcing effects mediated by dopamine-independent mechanisms in the nucleus accumbens⁵.

Other experimental approaches

The dopamine hypothesis has been a starting point for other investigations, ranging from molecular genetics to functional neuroimaging. For example, using positron-emission tomography, regional cerebral blood flow can be measured in humans after administering amphetamine-like drugs such as methylphenidate. Consistent with the results of animal studies, challenge with the drug leads to significant changes in the striatum, which correlate with subjective responses (such as euphoria and craving) in cocaine addicts⁶. Moreover, dopamine receptors have been investigated as products of possible 'candidate' genes to help explain why drug abuse tends to run in families⁷. One idea is that inherited variability in function of the dopamine D2 receptor explains why people differ in their responses to drugs of abuse, and why some run a greater risk of drug abuse or addiction.

Another big focus of interest has been the dopamine transporter. During transmission of a nerve impulse, dopamine neurons release this neurotransmitter into the synapse, which is a tiny gap between two neurons. Dopamine diffuses across the synapse and binds to receptors on the other side. Having done its job, dopamine is then recycled by the transporter, which facilitates re-uptake by the presynaptic neuron. But cocaine, amphetamine and methylphenidate all block such re-uptake by binding to the dopamine transporter, leading to increased levels of dopamine in the synapse (Fig. 2 on page 569). So, the transporter, as well as various dopamine receptors in the nucleus accumbens, has been a target for gene deletion or disruption ('knockouts') and other forms of genetic modification.

A graphic demonstration of the knockout approach is the finding that mice lacking the dopamine D2 receptor consume less alcohol than normal mice⁸. This result impli-

cates the dopamine system in the effects of ethanol, although ethanol affects many other neurotransmitter systems, including serotonin (also known as 5-hydroxytryptamine; 5-HT), the amino-acid transmitters GABA (γ -aminobutyric acid) and glutamate, and also neuropeptide Y. The knockout strategy has also produced results that do not readily harmonize with previous pharmacological evidence. For example, experiments on mice lacking the D2 dopamine receptor indicate that this receptor mediates some of the positive effects of morphine, but not the negative physical symptoms produced after withdrawal from the drug⁹. Yet this selectivity is contradicted by pharmacological evidence that the D2 receptor contributes to the physical dependence induced by withdrawal of morphine in rats, and that

the symptoms can be reduced by treatment with drugs that stimulate the D2 receptor¹⁰. Paradoxically, in mice lacking the dopamine transporter, the rewarding effects of cocaine are reduced — but not abolished — as measured by their tendency to self-administer the drug¹¹. This observation fits with the fact that, in human brain-imaging studies¹², subjective responses to cocaine do not correlate with its action at the dopamine transporter. And it is a challenge for the primacy of the dopamine hypothesis.

The serotonin challenge

The challenge to the dopamine hypothesis has recently been highlighted by the role of the serotonin neurotransmitter system in cocaine abuse. This is particularly relevant given the often mutually inhibitory interac-

tions between the serotonin and dopamine systems. Anti-depressant drugs such as Prozac share with cocaine a high affinity for the serotonin transporter molecule. Commonalities and co-morbidity between many forms of drug dependence and depression strengthen this link¹³. But again, the pharmacological and genetic evidence does not always seem to gel. For example, mice lacking the 5-HT_{1B} serotonin receptor are more likely than normal mice to self-administer cocaine¹⁴. Yet ostensibly the same behaviour can be achieved by the opposite action — treating rats with drugs that stimulate the 5-HT_{1B} receptor¹⁵.

Such controversies can be put down to the shortcomings of both approaches. For the pharmacological approach, the problem is one of selectivity — drugs are rarely selective for just one receptor. In the knockout mice, it's a case of functional compensation. The time lag between the genetic intervention and exposure of an adult mouse to cocaine means that, during development, there could be massive changes to compensate for the deletion of a particular receptor. The resulting 're-wired' brain may not function in the same way as normal, so invalidating — or, at least, greatly complicating — the use of knockout or transgenic animals to model adult drug dependence. We need not only more selective drugs, but also mice in which the affected genes can be regionally knocked out or induced; that is, synthesis of the protein in question can be turned off or on at a precise time in adulthood, preferably in specific regions of the brain.

Gene expression

Other molecular strategies may provide less ambiguous pointers to possible therapies. A complementary approach to that of deleting genes is to examine what happens to gene expression when particular drugs are repeatedly administered. This differs from the pharmacological and transgenic approaches in that it provides essentially correlative information; nonetheless, expressed gene products can be the subsequent targets of drug probes, with possible therapeutic dividends. There is growing information on the molecular maladaptations that occur once dopamine has bound to its receptors¹⁶. The two main classes of dopamine receptor (D1- and D2-like) can have opposite effects from one another on signalling pathways and gene transcription in postsynaptic neurons. In addition, chronic exposure to opiates, cocaine or ethanol decreases levels of some intracellular signalling molecules in the nucleus accumbens (inhibitory guanine-nucleotide-binding proteins, for example), but increases the activity of others (adenylyl cyclase and cyclic-AMP-dependent protein kinase), ultimately affecting gene transcription (Fig. 2).

Two of the best-studied families of tran-

Box 1 Common drugs of abuse

Psychomotor stimulant drugs This class includes amphetamines, cocaine and methylphenidate (which is used to treat hyperactive children). These drugs work directly on the monoamine (especially dopamine) neurotransmitter systems. They produce euphoria when taken intravenously, and the effects of withdrawal include dysphoria (mild depression), fatigue, sleep disturbances, increased appetite and anxiety. MDMA (methylenedioxymethamphetamine) or 'ecstasy' also falls into this class, although its main effects are probably on serotonin systems.

Opiates These analgesics, which include morphine and heroin (diacetylated morphine), are usually injected. Many of their subjective effects, including a sense of well-being and euphoria, act through opiate receptors of the μ type, to which naturally occurring chemical messengers such as β -endorphin and the enkephalins also bind. Symptoms of withdrawal include dysphoria, nausea, muscle cramps, tear

production, diarrhoea, sweating, anxiety and fever ('cold turkey').

Alcohol Acts in many ways, some of which are analogous to those of anxiety-relieving drugs such as the benzodiazepines (abuse of which can also lead to dependence). Withdrawal symptoms include autonomic hyperactivity, nausea, hand tremor, anxiety and hallucinations. Alcohol has striking sedative effects and can lead to memory loss.

Nicotine Works at receptors for the neurotransmitter acetylcholine, found, among other sites, in the neocortex, hippocampus and midbrain (Fig. 1). Withdrawal symptoms include dysphoria, insomnia, anxiety, restlessness, decreased heart rate and weight gain.

Cannabis Also known as hashish or marijuana, this drug is generally inhaled. It can produce a dependence syndrome, as well as mild cognitive impairment. The main active constituent is Δ^9 -tetrahydrocannabinol, which works at cannabinoid receptors in the hippocampal

formation, striatum and globus pallidus. A naturally occurring cannabinoid, anandamide, may be the normal occupant for these receptors.

LSD (lysergic acid diethylamide) Produces vivid hallucinations. Abuse occurs in a different pattern to that of other drugs. For example (with the possible exception of cannabis), LSD is the only drug that animals other than humans do not reliably self-administer.

Phencyclidine (PCP) Described as a 'dissociative anaesthetic', its effects include altered body image, feelings of isolation, cognitive disorganization and drowsiness, hostility and negativism, as well as euphoria and inebriation. It is more noted now as a model of human psychosis than as a major drug of abuse. High levels of the PCP receptor, which is part of the glutamate-*N*-methyl-D-aspartate receptor complex, are found in the hippocampus and neocortex, and intermediate levels occur in the amygdala, nucleus accumbens and caudate nucleus (striatum).

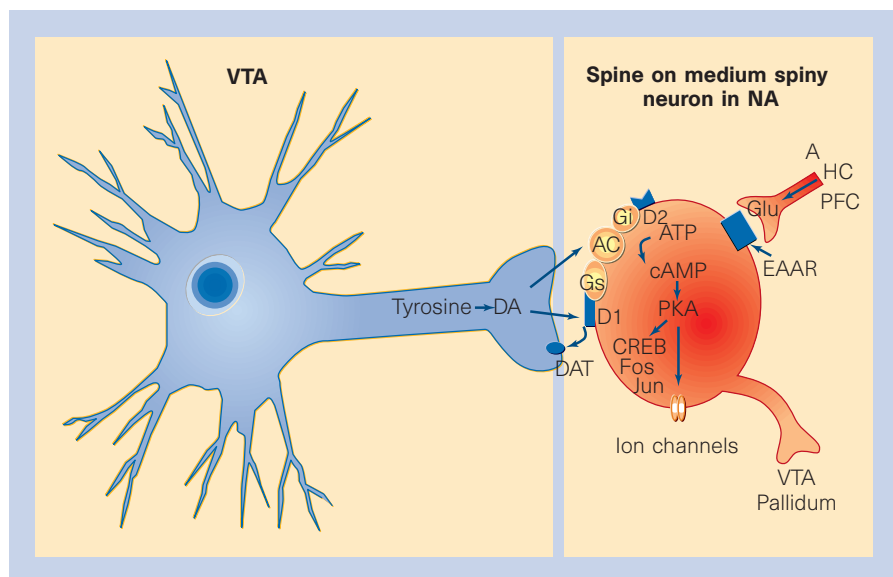


Figure 2 Neural systems of addiction. A dopamine-releasing neuron from the ventral tegmental area (VTA) is shown innervating a medium spiny neuron dendritic spine in the nucleus accumbens (NA). The dopamine transporter (DAT) is a main site for cocaine and amphetamine action. These drugs inhibit the re-uptake of dopamine by the VTA neuron, where it is initially produced from the amino acid tyrosine. Dopamine is shown acting at the two main families of dopamine receptor (D1 and D2). These are coupled to guanine-nucleotide-binding proteins (G_s and G_i), components of the intracellular cyclic AMP system, which also includes adenylyl cyclase (AC) and cAMP-dependent protein kinase (PKA). Possible substrates for this kinase include ion channels and the nuclear transcription factors CREB, Fos and Jun. A, amygdala; HC, hippocampus; PFC, prefrontal cortex; EAAR, excitatory amino-acid receptor; Glu, glutamate. (Adapted from ref. 16.)

scription factors are the so-called cAMP response element binding protein (CREB) family, and products of some immediate-early genes (referring to the time they are expressed after the ligand has bound to its receptor) such as *c-fos* and *c-jun*. Chronic exposure to morphine reduces levels of CREB. But chronic exposure to amphetamines may activate CREB in the nucleus accumbens, possibly by acting on the cAMP signalling pathway¹⁶. Certain Fos-like proteins (the Fos-related antigens) are also induced by chronic exposure to cocaine, amphetamine, morphine and nicotine, whereas mice that lack the *c-fos* gene show altered responses to cocaine¹⁶. Moreover, mice lacking the 5-HT_{1B} receptor show changes in their Fos-related antigens, even when never previously exposed to drugs¹⁴. This result indicates that the genetic make-up and subsequent development of these mice accelerates the molecular adaptation to drugs, correlating with a greater reinforcing effect.

Addiction as aberrant learning

All these discoveries support the consensus that drug dependence and addiction can be partly understood as gradual adaptations of the brain to chronic drug exposure. These adaptations may be triggered by a drive to regulate activity of the various brain systems within certain defined activity limits⁵, and they are the underlying processes for both the decreasing ('tolerance')¹⁷ and increasing ('sensitization')¹⁸ effects of repeatedly

administered drugs. They are consistent with the rebound consequences, following drug withdrawal, of chronic drug administration, which may further modulate the addiction process⁵. But how do the neurochemical and molecular changes produced by drug exposure relate to the clinical reality of human drug addiction?

First, we need to understand how addiction works at the cognitive, behavioural and neuropsychological levels. Several factors are increasingly seen as important for understanding variation in genetic and environmental vulnerability to drug abuse, and the treatment of addiction as a chronic, relapsing disorder. These include distinctions between the psychological, as well as neural, processes implicated in why people start to take drugs; 'consolidation' of these effects by the reinforcing action of the drug; subsequent maintenance of drug taking; and the eventual progression to addiction as a form of habit-based learning¹⁹. In the addicted stage, "the 'drug user' loses the voluntary ability to control its use" (ref. 19, page 237).

In the real world, drugs are not freely available, and the drug abuser has to forage for them. Drug-seeking behaviour can become powerfully associated with environmental cues, which, as 'conditioned stimuli', predict not only the availability of drugs (and their associated hedonic effects), but also aversive withdrawal states. The addict may seek to avoid such states by 'self-medication', through prophylactic or reactive drug taking.

The process of associative learning — by which the drug abuser connects specific cues such as a particular place with drug-induced states — probably includes structures in the brain that have strong anatomical links to the nucleus accumbens. These include the amygdala, hippocampus and orbitofrontal cortex, as well as related regions such as the pallidum and other sectors of the striatum (Fig. 1). Rats with damage to the amygdala readily self-administer cocaine, but cannot learn long sequences of behaviour to gain access to it²⁰. When human addicts are allowed to see the paraphernalia associated with giving themselves cocaine, several interconnected areas of the brain are activated — including the amygdala^{21,22}. As a result of such learning, behaviour is often sustained long after the original goal (in this case, cocaine) has been withdrawn. This may even lead to people developing a drug-seeking habit just as the subjective effects (for example, the 'rush' and euphoria) that initially encouraged it are reduced²³. If the addictive behaviour is, to some extent, ultimately divorced from the original drug effect that generated its development, 'surrogate' treatments, which mimic effects of the abused drug¹⁹ (such as D2-receptor agonists for cocaine abuse, methadone for opiate addicts, and nicotine patches for smokers), may have limited use.

The hypothesis that drug addiction is an aberrant form of learning, perhaps mediated by maladaptive recruitment of certain memory systems in the brain²⁴, is supported by several lines of evidence. First, receptors for both dopamine²⁵ and another neurotransmitter glutamate²⁶ are involved in normal learning in striatal and limbic structures. And second, transcription factors such as CREB are implicated in neuronal models of learning. Just as in other examples of learning, a conditioned stimulus alone (syringes, for example, or even a person associated with the drug) can activate the specific neural network that consolidated the original memory, through a series of plastic neuronal changes. In this way, the behaviour patterns associated with the drug and its attendant stimuli are evoked.

Relapse

Relapse, following what may often be protracted abstinence from drug taking, is a logical consequence of the conditioning process. It happens when the drug-seeking habit is reactivated by drug-related cues. Initially, the addict may retrieve from memory devastatingly compelling drug-related experiences (often described subjectively as 'craving'). This leads to further drug-seeking and drug-taking behaviour, sometimes even without the pleasurable anticipation associated with earlier drug experiences.

Relapse seems to be triggered by three main events²⁷. The first is taking the drug itself (although the effects of certain drugs,

Box 2 Common terms in psychopharmacology

Co-morbidity The statistically significant co-occurrence of distinct diseases or disorders within the same individual or group.

Conditioning Forms of learning by association. They may be either instrumental (in which voluntary actions are controlled by their outcomes; see reinforcers below) or Pavlovian (in which a temporal correlation between events is detected and learned by an animal, even in the absence of voluntary control).

Drug (substance) dependence A term

used synonymously with 'drug addiction'. It refers to the compulsive nature of drug seeking and taking, which precludes other forms of adaptive behaviour, impairing social and other forms of functioning. 'Physical dependence' refers to the need to take drugs to prevent aversive (usually involuntary 'autonomic') bodily symptoms caused by drug withdrawal.

Habit Technically, a form of instrumental learning in which a stimulus elicits a response without reference to the goal (or reinforcer) that originally motivated the learning. Aberrant 'habit'

learning may accurately describe the later stages of drug addiction, being a product of certain brain memory systems.

Homeostasis Classical physiological mechanisms of body regulation, which keep systems under equilibrium.

Reinforcement A reinforcer increases the likelihood of the act that produced it being repeated. The resultant strengthening of behaviour can be seen as a form of memory, by which voluntary actions with certain outcomes are consolidated into long-term memory.

such as morphine and cocaine, may substitute for one another). Second is a conditioned stimulus — for a rat or mouse, this could be a noise associated with the drug, which predicts its availability for self-administration. And the third is induction of a state of stress (although the pharmacological precipitation of withdrawal following opiate dependence does not seem to be effective). These three triggers are all thought to lead to similar neural events associated with drug taking (including the release of dopamine in the nucleus accumbens), which are enough to activate drug-seeking behaviour. Conditioning and memory retrieval may be good targets for treatment of addicts. For example, a drug that selectively stimulates another dopamine receptor (D3) greatly reduces cue-induced cocaine-seeking behaviour²⁸. And damage to the part of the amygdala that connects with the nucleus accumbens prevents cue-induced relapse for the same drug²⁹.

Inhibitory processes in the brain normally hold potentially maladaptive behaviour in check. Presumably, similar processes also restrain many of us from over-indulging on drugs in the first place. Such self-control is usually attributed to neural networks involving the prefrontal cortex and striatum (Fig. 1). In fact, some of the general behavioural and cognitive characteristics of drug abusers — including impulsivity (a tendency to act without foresight), risk taking and apparently poor decision-making abilities — resemble the effects of damage to the frontal lobes. For example, in decision-making tests, chronic amphetamine abusers perform similarly to patients with damage to the ventromedial (but not dorsolateral or medial) prefrontal cortex. A group of mainly opiate abusers, by

contrast, shows only part of this pattern of deficit³⁰. The similarity may be explained, in part, by findings that chronic amphetamine abusers have reduced serotonin function in the orbitofrontal cortex *post mortem*³¹. People who abuse 'ecstasy' show global reductions in binding of serotonin to the 5-HT transporter, based on measurements using positron-emission tomography³².

Correlations and causality

The observed correlations between drug taking and neuropsychological changes raise several points. First, the cognitive consequences of drug abuse and addiction may lead to problems of rehabilitation that far outreach just reducing drug-seeking behaviour. Second, the additional induced behavioural changes may accelerate the progression to addiction, for example, by impairing self-control. Third, the causal relationships between drug taking and neural, as well as neuropsychological, impairments are not clear. There may be a co-morbidity of drug-taking behaviour with other impulsive or risk-taking behavioural traits, because of genetic or developmental factors. Alternatively, drug taking could produce neural or neurotoxic 'side effects' that facilitate the drive to addiction. At present we cannot distinguish between these possibilities in humans although, for amphetamine users, the degree of the behavioural deficit on the decision-making task correlated with the length of time they had been abusing the drug³⁰.

Many questions remain. First, with drugs such as cocaine or amphetamine, does the dopamine system become less important when people become addicted, as control of behaviour devolves to other neural systems

that may mediate the habit-based learning? Second, how do we explain addiction to drugs that have distinct pharmacological actions? Most drugs of abuse have some common modes of action centred on the mesolimbic dopamine system, but they may also affect parts of the interactive memory systems that impinge on the striatum²⁴. Third, how do (possibly genetic) differences between people affect addiction — from the initial response to the drug to, for example, habit-formation mechanisms? Fourth, to what extent might chronic exposure to drugs lead to neurochemical changes that affect the habit-learning process? Finally, can molecular correlates of various phases of the addiction process be identified; for example, the 'switch' from 'drug-misuse' to addiction¹⁹? And what are the implications for the treatment of drug abuse, whether psychological or pharmacological, before and after this switch?

The examples given in this article illustrate the potential complexity of the factors that may influence addiction and its treatment. Drug-seeking habits can be consolidated and, in some cases, provoked or disinhibited by other effects of drug misuse. The remaining questions can be addressed at both the molecular and neuropsychological levels — it is to be hoped that the answers will lead to new treatments and therapies. □

Trevor W. Robbins and Barry J. Everitt are in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

e-mails: twr2@cus.cam.ac.uk

bje10@cus.cam.ac.uk

- Wise, R. A. & Bozarth, M. *Psychol. Rev.* **94**, 469–492 (1987).
- Hoebel, B. G. *et al. Psychopharmacology* **81**, 158–163 (1983).
- Bozarth, M. A. & Wise, R. A. *Life Sci.* **28**, 551–555 (1981).
- Rossetti, Z. L. *et al. Eur. J. Pharmacol.* **221**, 227–234 (1992).
- Koob, G. F. & Le Moal, M. *Science* **278**, 52–58 (1997).
- Volkow, N. D. *et al. Nature* **386**, 830–833 (1997).
- Uhl, G. *Neuropsychopharmacology* **20**, 3–9 (1999).
- Phillips, T. J. *et al. Nature Neurosci.* **1**, 610–615 (1998).
- Maldonado, R. *et al. Nature* **388**, 586–589 (1997).
- Harris, G. C. & Aston-Jones, G. *Nature* **371**, 155–157 (1994).
- Rocha, B. A. *et al. Nature* **393**, 175–178 (1998).
- Volkow, N. D. *et al. Nature* **386**, 827–830 (1997).
- Markou, A. *et al. Neuropsychopharmacology* **18**, 135–174 (1998).
- Rocha, B. A. *et al. Nature Neurosci.* **1**, 132–137 (1998).
- Parsons, L. *et al. J. Neurosci.* **18**, 10078–10089 (1998).
- Self, D. W. & Nestler, E. J. *Drug Alcohol Depend.* **51**, 49–60 (1998).
- Mendelson, J. H. *et al. Neuropsychopharmacology* **8**, 263–271 (1998).
- Robinson, T. E. & Berridge, K. C. *Brain Res. Rev.* **18**, 247–291 (1993).
- O'Brien, C. P. & McLennan, A. T. *Lancet* **347**, 237–240 (1996).
- Whitelaw, R. *et al. Psychopharmacology* **127**, 213–224 (1996).
- Grant, S. *et al. Proc. Natl Acad. Sci. USA* **93**, 12040–12045 (1996).
- Breiter, H. C. *et al. Neuron* **19**, 591–611 (1997).
- Tiffany, S. T. *Psychol. Rev.* **97**, 147–168 (1990).
- White, N. M. *Addiction* **91**, 921–949 (1996).
- Schultz, W. *et al. Science* **275**, 1593–1599 (1997).
- Herman, B. H. *et al. Neuropsychopharmacology* **13**, 269–293 (1995).
- Shaham, Y. *et al. J. Neurosci.* **16**, 1957–1963 (1996).
- Pilla, M. *et al. J. Psychopharm.* **11S**, A5 (1997).
- Meil, W. M. & See, R. E. *Behav. Brain Res.* **87**, 139–148 (1997).
- Rogers, R. D. *et al. Neuropsychopharmacology* **20**, 322–339 (1999).
- Wilson, J. M. *et al. Nature Med.* **2**, 699–703 (1996).
- McCann, U. D. *et al. Lancet* **352**, 1433–1437 (1998).

Myelin-like sheaths in copepod axons

Copepods, the small planktonic crustaceans that are the most abundant metazoans in the oceans, are so successful partly because they have an escape response that accelerates them to 200 body lengths per second within milliseconds^{1–3}. We find that nerve fibres of many copepods seem to be designed for rapid signalling. They have well-developed myelin-like sheaths, like those that give a tenfold boost to the conduction speed of nerve impulses in vertebrates⁴. By reducing a copepod's reaction time to predatory attack, these sheaths may be crucial to the survival of the large copepod populations that inhabit dangerous oceanic ecosystems.

Using transmission electron microscopy, we investigated the morphological substrates of this exceptionally quick reaction in copepods of the order Calanoida, which dominates planktonic communities and provides key links in marine food webs⁵. Cross-sections through the first antenna of 11 calanoid species revealed major contrasts in the way nerve fibres, or axons, are enveloped. Simple sheaths envelop axons in seven of the species, whereas multilayered sheaths resembling vertebrate myelin surround axons in the remaining four (Fig. 1a).

Species lacking the myelin-like sheaths belong to two superfamilies, the Arietelloidea and Centropagoidea. Although abundant, they are restricted in their distribution: the Arietelloidea are deep-sea forms, and the Centropagoidea occur primarily in freshwater and coastal marine habitats⁶. Species with myelinated sheaths belong to the Megacalanoida and Clausocalanoida, which are more widely distributed. These dominate open-ocean planktonic communities and also have representatives in coastal and deep-sea habitats.

In copepods possessing the myelin-like sheaths, nearly all axons of both sensory and motor nerves of the first antenna are ensheathed. Myelination of sensory axons is common among vertebrates but has not been reported previously among invertebrates⁷. Copepod myelin is characterized by concentrically arranged layers of membrane around the axon (Fig. 1b). The number of layers in a sheath varies for each axon, ranging from just one to more than fifty.

The highly organized laminae are tightly wrapped around microtubule-filled axonal cores. Bounding the core is a typical trilaminar unit membrane, or axolemma. The laminae consist of electron-dense layers of uniform thickness alternating with less dense layers of more variable thickness. The dark layers appear to represent a fusion of the apposed intracellular surfaces of sheath-cell membranes, without any intervening

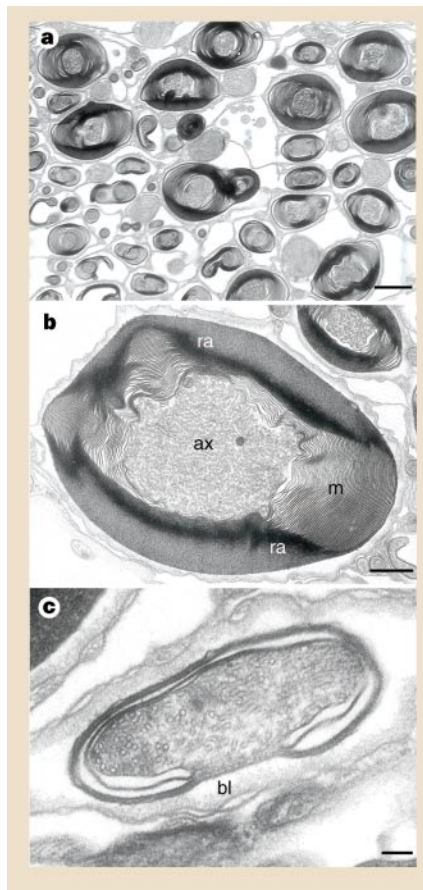


Figure 1 Transmission electron micrographs of cryo-fixed nerve in the first antenna of *Euchaeta rimana* (superfamily Clausocalanoida). **a**, Cross-section through sensory axons. All axons have myelin-like sheaths. Scale bar, 1 μm . **b**, Section of large motor axon, with many axonal microtubules, ensheathed by closely spaced myelin-like lamellae; ax, axon core; m, myelin-like laminae; ra, radial attachment zones. Scale bar, 0.5 μm . **c**, Focal node. Sheath laminae end against the axonal membrane, leaving the naked axon exposed to the basal lamina (bl). Scale bar, 0.1 μm .

cytoplasm, whereas sheaths in other invertebrates have a thin ribbon of cytoplasm between the membranous layers⁷. Sheaths in copepods have extensive radial attachment zones where the space between membranes is less than 1 nm (Fig. 1b).

In vertebrates, the myelin sheath is interrupted by circumferential regions called nodes of Ranvier, spaced along an axon to allow the flow of ionic current⁴. The myelin sheaths in copepods have discontinuities that resemble invertebrate patch or focal nodes⁷ (Fig. 1c), but none extends completely around the axon. The sheath laminae terminate at the axolemma in an orderly, stepwise fashion, with the innermost layers terminating first, allowing each layer to end in contact with the axolemma.

Considering its potential advantages, it is surprising that myelination is so rare in invertebrates⁷. Some do use myelin-like sheaths to increase the speed at which nerve impulses are conducted; indeed, speeds of 200 m s^{-1} in myelinated axons of penaeid shrimp are the highest in the animal kingdom⁸. The value of using neural specializations to decrease reaction times in small animals has been questioned because all parts of the animal are in very rapid communication anyway. But the existence of conduction-enhancing adaptations in small species, such as *Drosophila*⁹ and copepods¹⁰, and now including myelination in copepods, suggests a substantial value.

A preliminary comparison of copepod escape responses has shown significant differences in reaction times among species. *Undinula vulgaris*, which has myelinated axons, takes just 2 ms to initiate escape behaviour³; *Pleuromamma xiphioides*, which has only unmyelinated axons, takes 6 ms (unpublished data). In an animal with 2 mm of axon 5 μm in diameter between the antennal sensors and the central nervous system, and an estimated minimum of 1 mm more to the muscles, the conduction time for impulses travelling at 1.5 m s^{-1} in typical unmyelinated nerve¹¹ would be about 2 ms. For a tenfold increase in conduction speed due to myelination⁴, the same pathway would be 1.8 ms faster, explaining the faster reaction time of *U. vulgaris*. This improved performance may help calanoids with myelinated axons to dominate open-ocean communities.

April D. Davis*, **Tina M. Weatherby†**,
Daniel K. Hartline*, **Petra H. Lenz***

**Békésy Laboratory of Neurobiology, and †Biological Electron Microscopy Facility, Pacific Biomedical Research Center, University of Hawaii at Manoa, 1993 East-West Road, Honolulu, Hawaii 96822, USA
e-mail: petra@pbrc.hawaii.edu*

- Strickler, J. R. in *Swimming and Flying in Nature* Vol. 2 (eds Wu, T. Y.-T., Brokaw, C. J. & Brennan, C.) 599–613 (Plenum, New York, 1975).
- Lenz, P. H. & Hartline, D. K. *Mar. Biol.* **133**, 249–258 (1999).
- Trager, G., Achituv, Y. & Genin, A. *Mar. Biol.* **120**, 251–259 (1994).
- Ritchie, J. M. in *Myelin* (ed. Morell, P.) 117–145 (Plenum, New York, 1984).
- Mauchline, J. *Advances in Marine Biology* Vol. 33, *The Biology of Calanoid Copepods* (Academic, San Diego, 1998).
- Park, T. Phylogeny of calanoid copepods. *Synlogus* **58**, 191–196 (1986).
- Roots, B. I. in *The Node of Ranvier* (eds Zagoren, J. C. & Fedoroff, S.) 1–29 (Academic, Orlando, 1984).
- Kusano, K. *J. Cell. Physiol.* **68**, 361–384 (1966).
- Power, M. E. *J. Comp. Neurol.* **88**, 347–409 (1948).
- Yen, J., Lenz, P. H., Gassie, D. V. & Hartline, D. K. *J. Plankton Res.* **14**, 495–512 (1992).
- Bullock, T. H. & Horridge, G. A. *Structure and Function in the Nervous System of Invertebrates* Vol. 1 (Freeman, San Francisco, 1965).

Plant energetics and population density

Enquist *et al.*¹ present data from 37 plant species showing that the use of resources by individual plants scales approximately as the 3/4 power of plant mass, as predicted previously from a model of resource use in fractal-like branching structures². Thus, $Q \propto M^{3/4}$, where Q is the resource use, estimated as xylem transport, and M is the plant mass. In addition, their re-analysis of data from 251 populations¹ showed that the 'thinning law' between maximum plant population density ($N_{\max}$) and plant mass also obeys 3/4-power scaling, with $N_{\max} \propto M^{-3/4}$. From these two data sets, they inferred that population resource use per unit area is approximately independent of plant mass, with $N_{\max} Q \propto M^0$, a relation termed energy equivalence³.

Conversely (and more generally), the thinning law (with exponent $-y$, say) can be viewed as the consequence of energy equivalence and allometric scaling of individual resource use (with exponent y). This resource-based interpretation of the thinning law has been proposed previously for both plant⁴ and animal⁵ populations, but the theoretical derivation² of $y=3/4$ is an important new insight. Energy equivalence itself cannot be explained in any mechanistic sense by allometric scaling of individual resource use, despite apparent claims to the contrary¹, and remains to be accounted for as an empirical observation³.

I suggest that, for plant populations, energy equivalence reflects the facts that: (1) once plant canopies have reached closure, most of the incident radiation per unit area is intercepted^{6,7}; and (2) for well-watered plants, growth rate per unit of intercepted radiation (that is, the light utilization efficiency, or LUE) is approximately independent of plant mass⁸⁻¹⁰. There is now a mechanistic explanation for the latter observation in terms of leaf photosynthetic acclimation to light¹¹. It follows that, for closed canopies not subject to water limitation, population energy use for growth is roughly independent of plant mass, but may vary with incident radiation and LUE. An analogous argument, involving mass-independent population resource capture and utilization efficiency, might also explain energy equivalence in animal populations.

However, a word of caution is needed here. An important counter-example to energy equivalence is given by the well-documented observation that the growth rate per unit area of evenly aged forests eventually declines as individual trees become larger^{12,13}. Above-ground net primary productivity typically reaches a maximum in

young forest stands and then decreases by up to 76%, with an average reduction of 34% according to 13 studies of forest age sequences¹⁴. The rate of decline has important implications for sustainable forest management and the role of forests in the global carbon budget. This apparently universal phenomenon has been attributed, at least in part, to height-dependent hydraulic limitations on leaf stomatal conductance^{14,15}, implying that leaf photosynthetic rate and LUE may not always be independent of plant mass, particularly under water-limited conditions.

In summary, it is probably more appropriate to consider energy equivalence — like the thinning law — as an approximate rule of thumb, rather than as a fundamental law applicable to all plant types under all growth conditions.

Roderick C. Dewar

Unité de Bioclimatologie,
INRA Centre de Bordeaux,
BP 81, 33883 Villenave d'Ornon, France
e-mail: dewar@bordeaux.inra.fr

1. Enquist, B. J., Brown, J. H. & West, G. B. *Nature* **395**, 163–165 (1998).
2. West, G. B., Brown, J. H. & Enquist, B. J. *Science* **276**, 122–126 (1997).
3. Damuth, J. D. *Nature* **395**, 115–116 (1998).
4. Dewar, R. C. *Ann. Bot.* **71**, 147–159 (1993).
5. Lattin, J. *Oikos* **69**, 531–534 (1994).
6. Newton, I. E. & Blackman, G. E. *Ann. Bot.* **34**, 329–348 (1970).
7. Kozlowski, T. T., Kramer, P. J. & Pallardy, S. G. *The Physiological Ecology of Woody Plants* (Academic, San Diego, 1991).
8. Monteith, J. L. *Phil. Trans. R. Soc. Lond. B* **281**, 277–294 (1977).
9. Gallagher, J. N. & Biscoe, P. V. *J. Agric. Sci.* **91**, 47–60 (1978).
10. Cannell, M. G. R., Sheppard, L. J. & Milne, R. *Forestry* **61**, 123–136 (1988).
11. Dewar, R. C., Medlyn, B. E. & McMurtrie, R. E. *Plant Cell Environ.* **21**, 573–588 (1998).
12. Waring, R. H. & Schlesinger, W. H. *Forest Ecosystem Concepts and Management* (Academic, San Diego, 1985).
13. Ryan, M. G. & Waring, R. H. *Ecology* **73**, 2100–2108 (1992).
14. Gower, S. T., McMurtrie, R. E. & Murty, D. *Trends Ecol. Evol.* **11**, 378–382 (1996).
15. Ryan, M. G. & Yoder, B. J. *BioScience* **47**, 235–242 (1997).

Enquist *et al.*¹ present an interesting analysis of the link between plant size, allometry and mortality. But I believe that their claim that it has a functional basis is misleading when it is scaled up to whole populations.

Transpiration in plants is strongly driven by environmental conditions². It is therefore difficult to compare the transpiration rate per plant among different species unless these plants are exposed to the same environment. This does not apply here, as Enquist *et al.* derived their transpiration rates from a re-analysis of data from the literature. The confounding effect is negligible at the scale of individual plants, given the wide range of rates and dimensions reviewed. But when scaling up to the whole canopy, environmental driving variables such as radiation, water availability and site fertility predominate and tend to obscure any effects of plant size on the function of stands of species. This is clear from a comparison of maximum conductances and

assimilation rates among biomes across the world³.

The transpiration rates that have been scaled up to the level of populations by Enquist *et al.*¹ (see their Fig. 4) seem surprisingly high: rates of about 100 l m⁻² per day (that is, millimetres per day) far exceed the maximum values of 3–12 mm per day derived from a global comparison of plant canopies all over the world⁴. Another meta-analysis of evapotranspiration in coniferous forests and grasslands gave maximum values of 6–7 mm per day (ref. 5).

What is most important, however, is that the conclusions drawn by Enquist *et al.* from their upscaling to whole populations are misleading. It is debatable whether "total energy use or productivity of plants in ecosystems is... invariant with respect to body size": stand chronosequences in forest tree species⁶ indicate that, after canopy closure at the polestage, the leaf-area index tends to decline (as a result of self-thinning, among other processes). In mature canopies, this has a marginal effect on the interception of radiant energy and gross primary production⁷. Net assimilation and above-ground allocation, however, are further reduced by increasing respiratory costs, nutrient immobilization in soil litters and hydraulic constraints⁷⁻¹⁰, which are all a direct result of increasing body size, contributing to the well-known decline in forest growth with tree dimensions and age^{6,8}.

Even the conservative nature of forest evapotranspiration on the regional scale¹¹ seems to be true more at the community than at the population level, resulting from interaction among overstorey and understorey processes³. On the contrary, there are considerable changes in transpiration with stand development¹². The results of Enquist *et al.* do not seem to alter this picture.

Federico Magnani

Dipartimento di Produzione Vegetale,
Università della Basilicata,
Via N. Sauro 85, 85100 Potenza, Italy
e-mail: federico@imgpf.fi.cnr.it

1. Enquist, B. J., Brown, J. H. & West, G. B. *Nature* **395**, 163–165 (1998).
2. Monteith, J. L. & Unsworth, M. H. *Principles of Environmental Physics* (Arnold, London, 1990).
3. Schulze, E.-D., Kelliher, F. M., Körner, C., Lloyd, J. & Leuning, R. *Ann. Rev. Ecol. Syst.* **25**, 629–660 (1994).
4. Jones, H. G. *Plants and Microclimate* (Cambridge Univ. Press, 1992).
5. Kelliher, F. M., Leuning, R. & Schulze, E.-D. *Oecologia* **95**, 153–163 (1993).
6. Ryan, M. G., Binkley, D. & Fownes, J. H. *Adv. Ecol. Res.* **27**, 213–262 (1997).
7. Mencuccini, M. & Grace, J. *Tree Physiol.* **16**, 459–468 (1996).
8. Gower, S. T., McMurtrie, R. E. & Murty, D. *Trends Ecol. Evol.* **11**, 378–382 (1996).
9. Yoder, B. J., Ryan, M. G., Waring, R. H., Schoettle, A. W. & Kaufmann, M. R. *Forest. Sci.* **40**, 513–527 (1994).
10. Magnani, F., Borghetti, M. & Grace, J. in *System Analysis and Simulation for Agricultural Science: A Contribution of the RAISA Project* (eds Miglietta, F., van Laar, H. H. & Goudriaan, J.) (CABO-DLO, Wageningen, 1996).
11. Roberts, J. J. *J. Hydrol.* **66**, 133–141 (1983).
12. Rauner, J. L. in *Vegetation and the Atmosphere* (ed. Monteith, J. L.) 241–264 (Academic, London, 1976).

Enquist et al. reply — Energy equivalence, as originally defined¹, was an empirical relationship observed in animals: species differing in body mass, M , by many orders of magnitude tend to have almost equal rates of energy use per unit area by the population, because of an inverse allometric scaling relationship between energy use by the individual, or its metabolic rate, B , and the maximal population density, $N_{\max}$. Because $B \propto M^{3/4}$ and $N_{\max} \propto M^{-3/4}$, energy use is proportional to $BN_{\max} \propto M^{3/4}M^{-3/4} \propto M^0$.

We have shown that this also holds for plants², namely that the allometric scaling of both B and $N_{\max}$ appear to be the same as in animals. Further, we showed that the scaling of population density, and the resulting energy equivalence, can be explained by a simple model that assumes that plants grow until the allometric rates of resource use by the population equals the rate of resource supply.

In previous work, this was far from clear. The traditional explanation of the thinning law was not based on rates of individual resource use, but was thought to result from the geometric packing of canopies³ so that $N \propto M^{-2/3}$, or to be due to biomechanical constraints and the accumulation of non-living material⁴ so that $N \propto M^{-3/4}$. None of these studies, including that of Dewar⁵, has predicted the scaling of individual resource use. We not only make this explicit prediction, but also derive it from a model of resource distribution in plants with fractal-like branching structures⁶.

Dewar implies that energy equivalence reflects the fact that essentially all of the light is intercepted by closed canopies, and that growth rates of plants per unit of radiation are invariant. We agree, and in fact this is a special case of our general argument. Furthermore, we now have a detailed quantitative whole-plant model of resource distribution from which the general derivation of energy equivalence follows⁷. This model assumes that the size and photosynthetic rate of leaves are independent of plant mass. It predicts the number of leaves to be proportional to B , which is proportional to $M^{3/4}$. This is exactly the condition needed to give the density–mass relationship and therefore energy equivalence. But our model is much more general because it also applies to populations in which canopies are not closed and where water or nutrients, as well as light (as claimed by Dewar), are the critical limiting resources.

Dewar correctly observes that there may be a slight decrease in production or growth following canopy closure during secondary succession. But this is a point of detail. The hundred-fold variation in resource use, including any possible decline in production associated with age, is very small compared with the trillion-fold variation in vascular plant size. We emphasized that

ecosystems such as adjacent grasslands and forests, which have similar rates of resource supply but are dominated by plants differing by many orders of magnitude in mass, typically have similar rates of primary production. This does not seem to have been appreciated, despite work on relationships between population density and plant size and between individual plant performance and ecosystem productivity.

We agree with Magnani that transpiration and plant metabolism are influenced by local environmental conditions, especially when comparisons are made between individuals of relatively similar size^{8,9}. This accounts, for example, for the nearly two orders of magnitude variation in productivity of ecosystems dominated by similarly sized plants (our Fig. 4 in ref. 2); a good example would be to compare Arctic tundra with tropical grassland. Organism transpiration and metabolism, however, are even more strongly influenced by plant size (our Fig. 1). Thus, rates of xylem flux vary by only about two orders of magnitude as plant mass varies by 12 orders of magnitude from the smallest herbs and seedlings to the largest trees.

The rates of evaporation for ecosystems (our Fig. 4) are near-maximal short-term rates measured for individuals of known size of different plant species under near-optimal conditions. A few of these values are indeed higher than rates reported for ecosystems averaged over longer periods of months to years. Given this caveat, it is impressive that our estimates obtained by scaling up from individual plants (max, 82.4 l m⁻² per day; min, 1.0 l m⁻² per day; mean, 16.3 l m⁻² per day, s.d. 16.44) are within an order of magnitude of the maximum reported values. Most values fall within the range obtained using different sampling techniques, including remote sensing, for estimating the productivity of entire ecosystems¹⁰.

We therefore disagree with Magnani's objection to our extrapolation of the consequences of allometric scaling at the level of individual plants to processes that operate at the level of populations and ecosystems¹¹. Allometric relationships are not useful for understanding small differences in performance among similarly sized plants, such as forest trees following canopy closure. Allometry is invaluable, however, for understanding the pervasive effects of plant size on such diverse phenomena as the structure of single- and mixed-species stands, or on the productivity of ecosystems dominated by plants of contrasting size, such as adjacent agricultural fields, grasslands and forests. Understanding quarter-power allometric scaling relations at different levels of biological organization (in both plants and animals) will contribute to a common explanation of how body size

influences the acquisition and allocation of resources and thereby affects the abundance, distribution and diversity of sizes.

Brian J. Enquist*, **James H. Brown***, **Geoffrey B. West†**

The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

**Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA*

†Theoretical Division, T-8, MS B285, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

1. Damuth, J. *Nature* **290**, 699–700 (1981).
2. Enquist, B. J., Brown, J. H. & West, G. B. *Nature* **395**, 163–165 (1998).
3. Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. *J. Biol. Osaka City Univ.* **14**, 107–129 (1963).
4. Franco, M. & Kelly, C. K. *Proc. Natl Acad. Sci. USA* **95**, 7830–7835 (1998).
5. Dewar, R. C. *Ann. Bot.* **71**, 147–159 (1993).
6. West, G. B., Brown, J. H. & Enquist, B. J. *Science* **276**, 122–126 (1997).
7. Enquist, B. J., West, G. B. & Brown, J. H. in *Scaling in Biology* (eds Brown, J. H. & West, G. B.) (Oxford Univ. Press, in the press).
8. Kozlowski, T. T. & Pallardy, S. G. *Physiology of Woody Plants* (Academic, New York, 1997).
9. Lambers, H., Chapin, F. S. I. & Pons, T. L. *Plant Physiological Ecology* (Springer, New York, 1998).
10. Schulze, E. D., Kelliher, F. M., Körner, C., Lloyd, J. & Leuning, R. *Annu. Rev. Ecol. Syst.* **25**, 629–660 (1994).
11. Ehleringer, J. R. & Field, C. B. *Scaling Physiological Processes: Leaf to Globe* (Academic, New York, 1993).

Biostratigraphy of new pterosaurs from China

Pterosaurs are represented in China by five genera and some isolated bones ranging in age from the Middle Jurassic to the Late Cretaceous period^{1,2}. Four of these genera belong to the derived monophyletic subgroup Pterodactyloidea; only the Middle Jurassic *Angustinaripterus* from Dashanpu, Sichuan, is a non-pterodactyloid (traditionally 'rhamphorhynchoids', a paraphyletic taxon). Two further pterosaurs^{1,2} (Fig. 1) from the Chaomidianzi Formation of the Beipiao area, western Liaoning Province, occur in the Liaoning beds, several metres higher than the compsognathid coelurosaur *Sinosauropteryx* and the basal bird *Confuciusornis*. Our analysis of these two fossils and other components of the fauna suggest a Late Jurassic biostratigraphic age for the Liaoning beds, which are important in the study of avian origins.

One new pterosaur, *Dendrorhynchus curvidentatus*¹, is morphometrically most similar to the Late Jurassic (Tithonian) Solnhofen form *Rhamphorhynchus* but does not fit within that taxon. Principal components analysis (by K. I. Warheit and K. P.) (Fig. 2) of the long bones of various pterosaurs indicates that *Dendrorhynchus* clusters most closely with *Rhamphorhynchus* and has no obvious unique features, but its proportions differ from those



Figure 1 The new Liaoning pterosaurs. **a**, *Dendrorhynchoides* ("Dendrorhynchus") *curvidentatus* n. g. (Ji & Ji, 1998) (holotype, National Geological Museum of China GMV2128). Scale bar, 20 mm. **b**, Detail of *D. curvidentatus* skull. Scale bar, 20 mm. **c**, *Pterodactylus* ("Eosipterus") *yangi* Ji & Ji, 1997 (holotype, GMV2117). Scale bar, 50 mm.

of the ontogenetic trajectories of *Rhamphorhynchus*. Instead, like *Scaphognathus*, it is distinct at the generic level. The generic name *Dendrorhynchus* is preoccupied by a nemertine³, so we propose replacing it with the name *Dendrorhynchoides*.

The second new fossil, *Eosipterus yangi*² (Fig. 1c), is a large pterodactylid with a wingspan of about 1.25 metres. The incomplete fusion of the carpals and tarsals indicates that this specimen, although large for a pterodactylid, was not fully adult. A plot of the second and third principal components (the first was size) of the preserved long bones of *Eosipterus* (Fig. 2) groups *Eosipterus* with the Late Jurassic Solnhofen pterodactylids *Pterodactylus kochi*, *P. antiquus* and *Germanodactylus*, which have been viewed as part of a single ontogenetic series⁴ of *P. antiquus*, with which *Eosipterus* may be synonymous.

The age of the Liaoning beds has been argued to be Late Jurassic or Early Cretaceous, and radiometric dates give conflicting results⁵. The biostratigraphy of the fauna

through the entire section of the Yixian and Chaomidianzi formations is ambiguous. At Sihetun (in the Beipiao area), fish, frogs, turtles, lizards and mammals have been found, as well as theropod and sauropod (saurischian) and psittacosaurid (ornithischian) dinosaurs^{6,7}. Among the theropods, *Sinosauropteryx*, *Protarchaeopteryx*, *Caudipteryx* and *Confuciusornis*^{8–10} are of primary importance to studies of the origins of birds and avian features.

When exact faunal equivalents are not available, next-of-kin taxa can provide at least minimal divergence times. Psittacosaurid dinosaurs have so far been known only from the Early Cretaceous⁷, but the separation of these marginocephalians from other ornithischian dinosaurs was no later than Early Jurassic, perhaps much earlier¹¹. *Sinosauropteryx* is the most closely related of the Liaoning coelurosaurs to the Late Jurassic basal coelurosaur *Compsognathus* from the Solnhofen limestones of Germany⁸. *Protarchaeopteryx*³ is a basal maniraptoran, as are avians such as the Solnhofen form

Archaeopteryx, and the Liaoning form *Caudipteryx* has been linked basally to these taxa³. *Confuciusornis* is the next most basal bird known after *Archaeopteryx*. The coelurosaurian lineages therefore provide a biostratigraphic signal of sister-taxa rooted in the Late Jurassic. None of these specific genera is known from the Early Cretaceous, although related lineages persisted.

The two new Liaoning pterosaurs provide a similar signal by clustering with Late Jurassic relatives. Pterodactylids are known from both the Late Jurassic and the Cretaceous, but non-pterodactylid pterosaurs are not reliably known from any Cretaceous deposits¹². The Jurassic–Cretaceous transition among pterosaurs may therefore be sharper than in some other tetrapod faunal components, and the available data from coelurosaurs and particularly pterosaurs suggest a Late Jurassic age for the beds in which they are found. In contrast, the earliest Cretaceous faunas, such as that of the Wealden, bear little resemblance to the Liaoning and Solnhofen faunas¹³. But the most basal Cretaceous faunas must be better known before any biostratigraphic hypothesis can be considered iron-clad.

S.-A. Ji, Q. Ji, K. Padian*

National Geological Museum of China, Xisi, Beijing 100034, China

*Department of Integrative Biology and Museum of Paleontology, University of California, Berkeley, California 94720-3140, USA
e-mail: kpadian@socrates.berkeley.edu

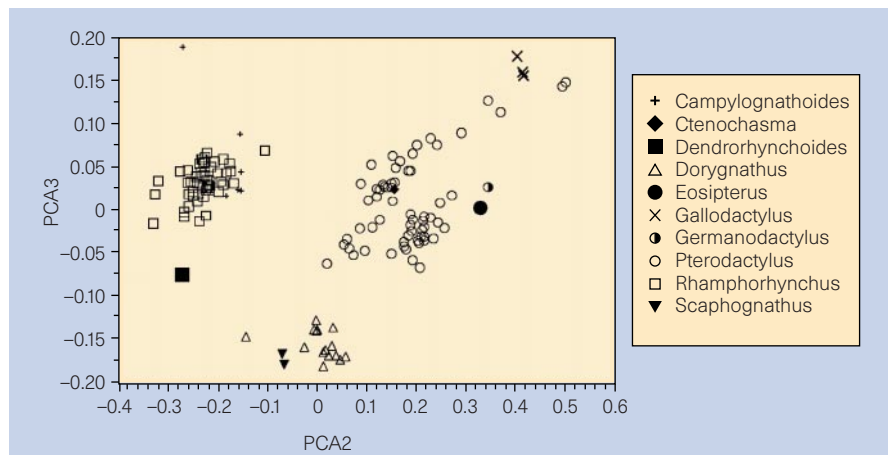


Figure 2 Principal components analysis (PCA). PCA plot of second and third major axes, comparing radius, wing metacarpal, first two wing phalanges, and tibia in *Campylognathoides*, *Ctenochasma*, *Dendrorhynchoides*, *Dorygnathus*, "Eosipterus"², *Gallodactylus*, *Germanodactylus*, *Pterodactylus*, *Rhamphorhynchus* and *Scaphognathus*. *Dendrorhynchoides* differs significantly from *Rhamphorhynchus* along the third major axis; "Eosipterus" falls within the cluster of points of the congeners *Pterodactylus*, *Germanodactylus* and *Ctenochasma*³ for pairwise comparisons of PCAs 1–3, and so is probably a large specimen of *Pterodactylus* (including *Germanodactylus*).

- Ji, S.-A. & Ji, Q. *Jiangsu Geol.* **22**, 199–206 (1998).
- Ji, S.-A. & Ji, Q. *Acta Geol. Sin.* **71**, 115–121 (1997).
- Yin, Z. & Zeng, F. *Oceanol. Limnol.* **16**, 323–335 (1995).
- Bennett, S. C. *J. Vert. Paleontol.* **16**, 432–444 (1996).
- Ji, Q. et al. *Prof. Pap. Stratigr. Paleontol.* (in the press).
- Wang, X.-L. et al. *Vert. Palasiat.* **36**, 81–101 (1998).
- Xu, X. & Wang, Y. *Vert. Palasiat.* **36**, 147–158 (1998).
- Chen, P.-J., Dong, Z.-M. & Zhen, S.-N. *Nature* **391**, 147–152 (1998).
- Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-A. *Nature* **393**, 753–761 (1998).
- Hou, L.-H., Martin, L. D., Zhou, Z.-H. & Feduccia, A. *Science* **274**, 1164–1167 (1996).
- Padian, K. in *Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 549 (Academic, San Diego, 1997).
- Padian, K. *Modern Geol.* **23**, 57–68 (1998).
- Norman, D. B. in *Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 783–785 (Academic, San Diego, 1997).

Prometheus unwound

Children of Prometheus: The Accelerating Pace of Human Evolution

by Christopher Wills
Perseus: 1998. 310 pp. \$25

Adrienne Zihlman

Books about human evolution tend to fall into two categories: they either emphasize how much like the apes we are in terms of our murderous, hierarchical or sexual behaviour, or they elaborate the contrast between hairy, jungle-bound quadrupeds and our brainy upright selves. Christopher Wills, a molecular researcher, takes the second tack. In his view, we children of Prometheus have stolen fire, language and high technology from the gods of things-as-they-are and have evolved, and we continue to evolve so rapidly that we have left our relatives, the chimpanzees, gorillas and orang-utans, in the Darwinian dust.

Although Wills acknowledges that human and chimpanzee DNA are 98% identical, he is even more impressed by the degree to which "our big-brained, clever-handed, highly social, language-speaking selves" have outpaced our hirsute relatives in the hominoid Olympiad.

In support of his thesis that the rate of human evolutionary change is accelerating, Wills recounts a number of case studies: how the Tibetans have acquired physiological adaptations to high altitudes; how human haemoglobin variants confer resistance to malaria; and even how one's position in the hierarchy of the British Civil Service affects longevity. He profiles the changing character of Europe, with declining birth rates and rising immigration that will mix gene pools and lead to increased diversity.

Turning from the human present to its past, Wills reviews the latest versions of the human fossil record and concludes that the story of the extinct and presumably less successful Neanderthals represents "the road not taken" by our own species. This judgement may be a trifle premature, as the Neanderthals inhabited Europe and the Middle East for several hundred thousand years and *Homo sapiens* has been around for no more than 200,000 to date. The extensive morphological differences between Neanderthals and modern humans is barely mentioned, but the DNA comparison reported last year is recounted in entertaining detail.

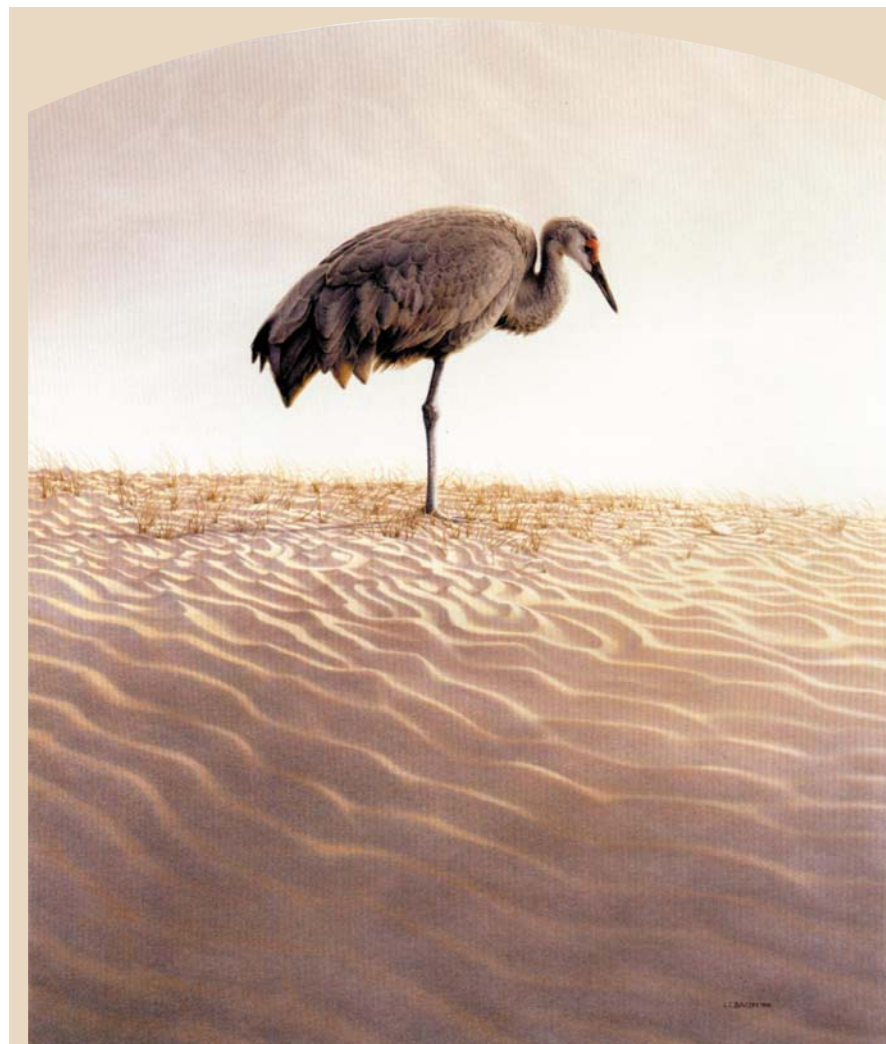
Examining the earlier fossil record with an eye to 'molecular-clock' timing of the human-chimpanzee divergence to about five million years ago, Wills 'predicts,' as others have done, that the earliest hominids will be chimp-like, and in fact australopithecines very much resemble bipedal chimpanzees.

Wills unquestioningly accepts Tim White's claim that his 4.4-million-year-old *Ardipithecus ramidus* is the earliest-known hominid, although the lack of published evidence for the fossil's bipedality has aroused considerable scepticism in the palaeo-anthropological community. Wills does not mention Meave Leakey's more recent discovery in Kenya of *Australopithecus anamensis*, an undisputed hominid possessing limb bones that demonstrate bipedality four million years ago.

This book, though often interesting and informative, seems to suffer from imprecision in the use of the word 'evolution' and from a lack of coherent focus. It is not always clear whether Wills is referring to genetic, morphological or cultural evolution, as he

switches vertiginously from one to another. Sometimes his argument rests entirely at the molecular-genetic level (changes in gene frequencies in populations), and the morphological level is omitted, as in the Neanderthal account. At other times, he abandons quantitative molecular data and builds a hominoid family tree from a motley array of physical and behavioural features, "including the ability to speak, the ability to recognize oneself in a mirror, the ability to walk bipedally, the degree of difference in size between the two halves of the brain, the amount of thumb mobility, the period over which the young are nursed, the amount of skeletal muscle strength, the presence of breast in a non-lactating female, and a number of others".

Wills concedes that chimpanzees are



The meaning of wildlife

Nineteenth-century artists such as Landseer tend to be sentimental or sensational in their representation of animals. The 89 artists whose work is shown in *Modern Wildlife Painting* by

Nicholas Hammond (Yale University Press, \$50, £35) aim to avoid the anthropocentric. Still, Chris Bacon admits this painting of a sandhill crane symbolizes his own musings about life.

stronger than humans, but deduces from his anthropocentric items that human evolution has occurred at about 10 times the rate of ape and monkey evolution.

How easy it would be to argue the contrary from the chimpanzee point of view: that we are a most retrograde species who have lost our body hair, facility in climbing trees, female sexual swellings and numerous other primate assets. A more neutral description would merely state that we are adapted to a way of life very different from the ancestral one, and that small changes at the molecular-genetic level can obviously result in major phenotypic and behavioural modifications.

The mystery of the big human brain and what goes on inside it intrigues Wills, as it has fascinated so many others.

The last few chapters, some of the best in this book, deal with the genetic aspects of diseases such as narcolepsy, attention-deficit hyperactivity disorder, schizophrenia, Alzheimer's disease and the much-debated factors in IQ. He speculates that rapid generational changes in microsatellite DNA — which are known to account for several neurological diseases, such as Huntington's chorea — could be a mechanism for intellectual improvement. In Wills's opinion, our grandchildren will surely be smarter than us, and within a few generations Einsteins and Mozarts will be commonplace.

These optimistic exegeses, while not without interest, illuminate by omission how little we understand about the molecular, functional and adaptive factors that have

inflated the human cranium and its contents threefold over the past five million years.

In his last chapter, "Our evolutionary future", Wills unhesitatingly predicts an even bigger-brained advent that will carry us beyond the Solar System to planets of other stars. I admire the courage of someone who can make such projections for a society that cannot predict the next election result or the weather next week.

Are we still evolving? This is the question Wills asks at the beginning and end of his volume, and at many points in between. The answer is, of course, that we are: as are chimpanzees, gorillas, orang-utans, crocodiles and bacteria. Are we really evolving much faster than those other organisms? Not at the molecular level, according to present evidence. Even bacteria, the oldest and most 'primitive' forms of life, show a degree of flexibility in their ability to survive and become resistant to our most potent antibiotics that our own species has yet to prove it can match.

In Darwin's concept of evolution, the direction of evolutionary change is unpredictable, dependent as it is on changing conditions and the vagaries of natural selection. Wills's confidence as to which evolutionary features are good or bad, superior or inferior, faster or slower, conveys perhaps a hint of Promethean hubris.

For the future of the species, let's hope that our grandchildren will, as Wills predicts, be smarter than we are. □

Adrienne Zihlman is in the Department of Anthropology, University of California, Santa Cruz, California 95064, USA.

A most uplifting career

Out of the Crater: Chronicles of a Volcanologist

by Richard V. Fisher

Princeton University Press; 1999. 178 pp.

\$24.95, £16.95

Steve Sparks

The life of an academic field geologist, especially one interested in volcanoes, has considerable attractions. Pleasures to be found include visiting beautiful and sometimes remote parts of the world, and enjoying the companionship of like-minded colleagues, the inevitable adventures and mishaps, and the intellectual challenges of interpreting the tantalizing clues nature provides on how the Earth works. Richard Fisher, professor of geology at the University of California, Santa Barbara, is among the foremost volcanic geologists of the second half of this century. *Out of the Crater* is essentially his reminiscences.

The book makes interesting and entertaining reading because of the richness of Fisher's experiences and the spectacular nature of his subject. Each short chapter includes anecdotes and descriptions of the parts of the world he has visited, occasional explanations about the science of volcanology and some insights into how he developed his scientific ideas. The reader will be taken to some of the world's great volcanoes — such as El Chichón in Mexico, Vesuvius and Mount St Helens — and to diverse cultures: the American West, Argentina and China.

Fisher's unique formative experience was within the context of a remarkable episode in twentieth-century history. In 1946, at 17 years old, he was stationed in the army at Los Alamos and volunteered to go to Bikini atoll, where he observed the atomic bomb tests. His descriptions of the Bikini tests are fascinating, and highlight the military authorities' almost complete lack of understanding of the lethal and long-term effects of radioactivity. To swim in the ocean at Bikini atoll a few days after the test now seems incredible.

What made the Bikini experience so important for Fisher's later scientific achievements was seeing the base surge phenomenon — the cloud of radioactive dust and water droplets that spreads radially along the ground from the base of an atomic explosion. Fisher's significant contribution to volcanological science has been the elucidation of the geology of pyroclastic density currents, a term that encompasses volcanic base surges, pyroclastic flows and volcanic blasts like that of Mount St Helens in 1980.

When volcanic base surges were first recognized, Fisher realized the significance immediately. The enigmatic sedimentary structures in many pyroclastic sequences,

Blast from the past

This satellite image shows an unusual view of Mount St Helens and the area around it, 11 years after the 1980 eruption. Winds of up to 440 kph flattened a wide area and lava fanned out from the volcano, leaving what is now a horseshoe-shaped crater, lower left, containing a lava dome 300 metres high. Faint traces of green in the purple lava show the regrowth of vegetation, especially in protected valleys. It is one of the many dramatic computer-generated digital images, obtained by satellite sensing systems, in *America from Space* by Thomas B. Allen (Firefly, \$29.95).



New in paperback

Cats' Paws and Catapults: Mechanical Worlds of Nature and People

by Steven Vogel
Penguin, £8.99

"Vogel's book has been written to be enjoyed. ... It is full of ideas and well-explained principles that will bring new understanding of everyday things both to non-scientists and scientists alike". R. McNeill Alexander, *Nature* 392, 881 (1998).

The Number Sense: How The Mind Creates Mathematics

by Stanislas Dehaene
Penguin, £8.99

"Our ability to use numbers is as much a fundamental aspect of what makes us human as is our ability to use language. This book provides an excellent introduction to how this ability is organized in our mathematical brain." Brian Butterworth, *Nature* 391, 856 (1998).

Volcanoes: Crucibles of Change

by Richard V. Fisher, Grant Heiken and Jeffrey B. Hulen
Princeton University Press, \$19.95, £14.95

Twins And What They Tell Us About Who We Are

by Lawrence Wright
Wiley, £14.95

"As a measured analysis of a complex and controversial field, it is not very good... The book

is essentially a collection of the sometimes divergent views of the field's main players about the significance of genes." John Galloway, *Nature* 392, 34-35 (1998).

Why People Believe Weird Things: Pseudoscience, Superstition and Other Confusions of Our Time

by Michael Shermer, with a foreword by Stephen Jay Gould

W. H. Freeman, £10.95, \$14.95

"Shermer refuses to engage deeply with why beliefs are held. ... In the Middle Ages it took the rack and the thumbscrew to make people confess that their father drank babies' blood. Now, a couple of sessions in an air-conditioned psychotherapist's office seems to suffice. What needs explaining is how we have become such wimps." John C. Marshall, *Nature* 389, 29 (1997).

Rethinking Visual Anthropology

edited by Marcus Banks and Howard Morphy
Yale University Press, \$18, £12.95

Space and The American Imagination

by Howard E. McCurdy
Smithsonian Institution Press, \$17.95, £10.95

"McCurdy, an acknowledged authority on the US space programme ... tells a richly detailed and persuasively balanced story that explains the paradox of manned spaceflight." Alex Roland, *Nature* 392, 143-145 (1998).

which were commonly dismissed as reworking, were primary features of volcanic flow.

Fisher generously acknowledges that the credit for seeing the link between the atomic base surges and volcanic phenomena belongs to Adrian Richards, who briefly noted the similarities in 1959, and to Jim Moore of the US Geological Survey, who developed the insight in detail. However, it was principally Fisher who pioneered the systematic documentation and interpretation of the deposits of base surges with colleagues and students such as Aaron Waters, Hans Schmincke and Grant Heiken. This exciting and important period in volcanic geology is described with the insights into the twists and turns of original discovery that are usually missing from scientific papers.

Later in the book, Fisher also describes in characteristically modest fashion his important discoveries and ideas about pyroclastic processes such as ash-cloud surges, flow transformations and the expanded character of some pyroclastic flows.

Volcanology can be a dangerous profession. The most poignant part of the book concerns the deaths of two outstanding young American scientists by volcanoes just as their careers were blossoming. When Mount St Helens started to erupt in 1980, Harry Glicken was stationed at the Coldwater II observation post, six miles north of the volcano. Glicken

had a meeting with Fisher on 18 May to discuss his career and so David Johnston took over his duty. The catastrophic eruption of that day killed Johnston. Eleven years later, Glicken was killed by a pyroclastic flow on Mount Unzen in Japan along with the film-makers Katia and Maurice Kraft and many journalists.

Pyroclastic density currents are the most lethal of volcanic phenomena. As this book reminds the reader, advances in understanding the geology of these deposits (in which Fisher himself has played a leading role) have established the criteria for identifying them — which is an essential part of volcanic hazard and mitigation work.

Fisher is a good writer and mixes tales of his travels, science and personal anecdotes well. The book is entertaining and an easy read. Geologists will enjoy the book because they will be reminded of their own escapades. Young people contemplating a geological career will surely also be seduced by the idea of a slightly anarchic life in remote and beautiful parts of the world. Fisher has clearly been motivated by the wonder of nature and the fundamental desire to understand its mysteries that drives many successful natural scientists. Perhaps the only regret is that the book is not a bit longer. □

Steve Sparks is at the Chiron Corporation, D-207Q, Emeryville, California 94608, USA.

Speaking in tongues

The Neurolinguistics of Bilingualism: An Introduction

by Franco Fabbro
Taylor & Francis: 1999. 255 pp. £39.95
Risto Näätänen and Teija Kujala

The past few decades have seen our exposure to multiple languages continuously increasing. In this sense, Franco Fabbro's *Neurolinguistics of Bilingualism*, dealing with the representation and processing of multiple languages in the brain, meets an urgent need. Another major factor underlining the importance of this research field is that bilingualism is by no means unusual. If we define it in the broad sense — to include not only people who from an early age have spoken two or more languages, but also those who master a foreign language to a certain extent or a language in more than one dialect — then bilingualism covers more than half of the world's population.

The main message of the book is how lesions in various parts of the brain affect the comprehension and production of languages. Important questions are addressed, such as what happens to the native language compared with those acquired later in life when certain parts of the brain are damaged, or, if these languages are lost, which of them recovers first or to the greatest extent. The cases described elucidate not only bilingualism, but also, more generally, the fascinating complexity of human cognitive brain functions.

A great strength of the book is that it is organized in an extremely reader-friendly way. It starts with a general introduction on what language is and how the acoustical analysis and production of speech take place. This is followed by an introduction to the functional anatomy of the brain. The reader is also briefly introduced to technologies and means of studying cognitive brain functions and assessing language-related dysfunctions. This introductory part makes the rest of the book comprehensible to readers who do not have a solid knowledge of the brain's functional neuroanatomy. Without such an introduction, the main part of the book, which attempts to explain bilingualism from observations on patients with brain lesions causing various forms of aphasia and dysphasia, might be hard for readers outside the field of neuroscience to understand.

The case reports — describing how language functions can be affected by a brain lesion, and how these functions are recovered — are interesting and very helpful for understanding how multiple languages are represented in the brain. But this emphasis on brain lesions makes the book somewhat one-sided and is problematic for several reasons.

For example, if a region of the brain is damaged, causing loss of certain functions, it is hard to tell whether these functions are represented in that specific area or whether the loss is caused by a disconnection of two areas whose interaction is required.

Very recently, brain-activity measurements have shed new light on language representation and learning in the healthy human brain. For example, by using a cortical brain response termed 'mismatch negativity', it is possible to study the long-term memory representations of speech sounds of an individual's native language. Subsequent studies have shown how these native-language representations develop in early childhood, and what plastic changes accompany the development of the new phonetic categories of a foreign language learned in adulthood. Since this is a fairly recent development in neurolinguistics, however, it is understandable that it is not covered in the book.

In conclusion, this interesting and very clearly written book provides a good introduction to the way in which multiple languages are represented in the brain based on what we know from brain-lesioned patients. After reading it, an enthusiastic neurolinguist or neuroscientist will look forward to the next volume, elucidating bilingualism from studies on the healthy human brain. □

Risto Näätänen and Teija Kujala are at the Cognitive Brain Research Unit, Department of Psychology, PO Box 13, FIN-00014 University of Helsinki, Helsinki, Finland.

Hallo capitalism, goodbye science

What Have We Learned about Science and Technology from the Russian Experience?

by Loren R. Graham
Stanford University Press: 1998. 165 pp.
\$14.95, £10.95 (pbk).

Zhores A. Medvedev

The early history of Soviet science was very tragic indeed. Thousands of scientists, both prominent and unknown, perished in the Gulag. Biologists were forced to follow the pseudoscientific theories of Lysenko and other charlatans. The best aeroplane designers had their engineering facilities attached to special prisons. Promotion to a senior position, even in universities or institutes engaged in fundamental research, required membership of the Communist Party. Foreign travel was impossible. In libraries, even *Nature* fell prey to the censor's scissors.

But despite all this, Soviet science behind the Iron Curtain did reasonably well. It gave Stalin atomic and thermonuclear bombs much more quickly than did the scientists working for the democratic governments of

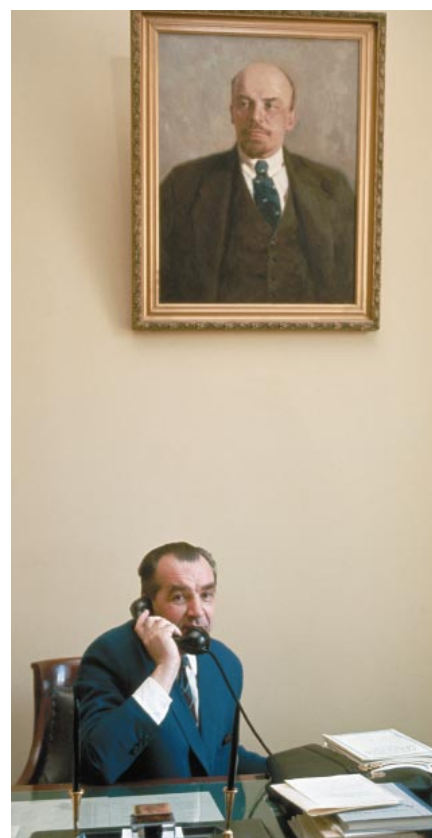
Britain and France, and it made antibiotics available for free health care throughout the USSR. The volume of research, basic and applied, and the number of research institutes grew rapidly. With Sputnik a few years later Soviet science and technology started to receive great respect and attention. In 1960, when Loren R. Graham arrived at Moscow University to study the history of the Academy of Sciences of the USSR, the Soviet Union had more researchers and engineers than the United States, and there were many areas of excellence, innovation and technological breakthroughs. Soviet scientists were world leaders in several fields.

In later years, Soviet science and technology continued to grow rapidly, but its efficiency did not keep pace with its size. This was mainly due to poor cooperation among the main sectors of research activity (universities, academies and ministries) and the vast military complex that consumed so much of the research and development budget.

Why, in new and democratic Russia, were science and technology treated so differently and deprived of financial support? Why did the transition from totalitarianism to democracy, from socialism to capitalism, harm the research capabilities much more than the infrastructure of science itself?

Graham, a leading historian of Soviet science and technology, tries his best to avoid political analysis and to link the developments in science with the general reform process. In his earlier books, Graham, like many other American historians, did not predict that science would collapse together with communism. It was believed that the development of science and technology in communist countries had a 'westernizing' influence, and that "science and technology have helped to make Soviet society more like the rest of the world, eroding the revolutionary and exceptional ethos in which the USSR was born" (*Science and the Soviet Social Order*, Harvard University Press, 1990). Among Soviet scientists, there was a general expectation that being free to exchange information, choose their collaborators and travel would aid their research, increase productivity and promote their integration into the world of science.

There was nothing wrong with this theory, or its expectations. The failure was with the transition itself. It was initiated as a step-by-step process by Mikhail Gorbachev, but was accelerated into the collapse of the USSR by Boris Yeltsin. As a result, a much smaller and poorer Russia was left with a research infrastructure that was too large and expensive, and most of the problems of 'big science' and military-industrial development became entirely irrelevant. With the largest external debt in the world, not only science and technology, but all state-dependent services — health, education, security and police — have been forced onto a survival regime with



Picture of health: Nikolai Blokhin, president of the USSR Academy of Medical Sciences, in 1964.

a subsistence level of government support. The current Russian budget is approximately the same as Denmark's.

Graham describes how some famous Russian research institutes have responded to poverty. They have usually reduced research expenses while trying hard to avoid redundancies. When resources get too low, they reduce salaries, often to levels of destitution, but try to keep staff, even if scientists cannot do much, or anything. They let parts of their — fortunately large — buildings to foreign companies or commercial firms and transform their fenced territories into paid parking lots for the expensive cars of the new elite.

This is not only a survival technique, but humanitarian as well. There are no redundancy payments, early retirement schemes or even unemployment benefits in Russia for scientists who lose their jobs. Some other state services, such as health, do the same. People continue to work even when they are not paid for months. They know that if they stop the situation will get worse, not better. But if we really want to take something from the Russian experience, it is that Russia is doing much better than the other republics of the former Soviet Union. Scientists in the Ukraine have started to wonder, half-jokingly, if their government will introduce entrance fees for their institutes. □

Zhores A. Medvedev is at the National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK.

Structural basis for recognition of the *tra* mRNA precursor by the Sex-lethal protein

Noriko Handa*, Osamu Nureki*†, Kazuki Kurimoto*, Insil Kim*, Hiroshi Sakamoto‡, Yoshiro Shimura§||, Yutaka Muto* & Shigeyuki Yokoyama*†

* Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

† Genomic Sciences Center and Cellular Signaling Laboratory, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-0106, Japan

‡ Department of Biology, Faculty of Science, Kobe University, 1-1 Rokkodaicho, Nada-ku, Kobe 657-0013, Japan

§ Department of Biophysics, Faculty of Science, Kyoto University, Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606-8224, Japan

|| Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

The Sex-lethal (Sxl) protein of *Drosophila melanogaster* regulates alternative splicing of the transformer (*tra*) messenger RNA precursor by binding to the *tra* polypyrimidine tract during the sex-determination process. The crystal structure has now been determined at 2.6 Å resolution of the complex formed between two tandemly arranged RNA-binding domains of the Sxl protein and a 12-nucleotide, single-stranded RNA derived from the *tra* polypyrimidine tract. The two RNA-binding domains have their β -sheet platforms facing each other to form a V-shaped cleft. The RNA is characteristically extended and bound in this cleft, where the UGUUUUUU sequence is specifically recognized by the protein. This structure offers the first insight, to our knowledge, into how a protein binds specifically to a cognate RNA without any intramolecular base-pairing.

In eukaryotic cells, mRNA precursors (pre-mRNAs) undergo a series of post-transcriptional events, such as processing, splicing, translocation to the cytoplasm, translation and degradation, in which numerous RNA-binding proteins are involved¹. The *Sex-lethal* (*Sxl*) gene product is an RNA-binding protein that plays a key role in sex determination and dosage compensation in *Drosophila melanogaster*². The Sxl protein induces female-specific alternative splicing of the *transformer* (*tra*) pre-mRNA^{3,4}. Sxl binds tightly to a characteristic uridine-rich polypyrimidine tract at the non-sex-specific 3' splice site in one of the *tra* introns^{3,4}, preventing the general splicing factor U2AF from binding to this site and forcing it to bind to the female-specific 3' splice site⁵. Sxl binds also to its own pre-mRNA and promotes female-specific alternative splicing^{6,7}.

The Sxl protein consists of 354 amino-acid residues and has two ribonucleoprotein (RNP) domains². The RNP domain, also called the RNA-recognition motif (RRM) domain and the RNA-binding domain (RBD), is one of the most common RNA-binding modules, and is thought to mediate RNA recognition in hundreds of proteins involved in post-transcriptional processes^{8–10}. The RBD consists of ~80 amino-acid residues with two highly conserved short motifs, RNP1 and RNP2 (refs 8–10) (Fig. 1). The crystal structure of the amino-terminal RBD of the U1 small nuclear (sn)RNP A protein (U1A) revealed that the RBD forms a four-stranded antiparallel β -sheet packed against two α -helices, with the RNP1 and RNP2 motifs in the two central β -strands¹¹. Tertiary structures of the U1A RBD–RNA complex have been reported: the crystal structure of the complex with the U1 snRNA hairpin II (ref. 12), and the NMR structure of the complex with an internal-loop RNA derived from the U1A pre-mRNA¹³. The crystal structure has been solved for a U1A-like protein, U2B', in a ternary complex with the U2A' protein and the U2 snRNA hairpin IV (ref. 14). These target RNAs have intramolecular base pairings that form stem/loop secondary structures, which are essential for the RNA–protein recognition.

On the other hand, the Sxl tandem RBDs are typical of the RBDs that have RNA-recognition mechanisms that differ from that of the U1A RBD. Various RBD-containing proteins bind to single-

stranded RNA with no base pairs¹⁰. The Sxl-binding polypyrimidine tract of the *tra* pre-mRNA is 5'-UUUUUGUUGUUUUUUUU (ref. 4), which seems to have no base pairs. An Sxl fragment consisting of the two RBDs connected by a ten-residue peptide retains nearly the same RNA-binding activity and specificity as the full-length protein^{15–17}. As a high proportion of RBD-containing proteins have multiple RBDs that are indispensable for their proper functions, the Sxl protein needs the two RBDs for site-specific binding to RNA. The isolated Sxl RBDs show much weaker affinity and specificity towards the target RNA^{15–17}. The solution structures of the individual Sxl RBDs, RBD1 (the first or N-terminal RBD)¹⁸ and RBD2 (the second or carboxy-terminal RBD)¹⁹, have the same global fold found in U1A and other RBDs. The RNP2 motif of the Sxl RBD1 is particularly different from those of other RBDs (ref. 18). The amino-acid sequence of the RBD1–RBD2 region, including the characteristic RNP2 motif of RBD1, of the Sxl protein is well conserved in the ELAV family of proteins¹⁸.

Here we present the crystal structure of the complex between the two tandemly arranged RBDs of the Sxl protein and a 12-nucleotide single-stranded RNA derived from the *tra* polypyrimidine tract. The structure reveals how the RBDs cooperate specifically to recognize the sequence, which has no intramolecular base-pairing, through extensive interactions with both the bases and the backbone moieties.

Structure determination

Crystallization trials were carried out for combinations of various Sxl fragments containing the RBD1–RBD2 region and various synthetic RNAs of different lengths and sequences. Crystallization trials with the wild-type Sxl fragments were unsuccessful, owing to the low solubility of the protein. The best crystals were obtained with a fragment encompassing residues 122 to 294 and carrying one mutation, Phe 166 → Tyr (Fig. 1), and a 17-nucleotide RNA, 5'-UUUUUGUUGUUUUUUUU, which was derived from the uridine-rich polypyrimidine tract before the regulated 3' splice site of the *tra* pre-mRNA. This mutant protein shows nucleotide-

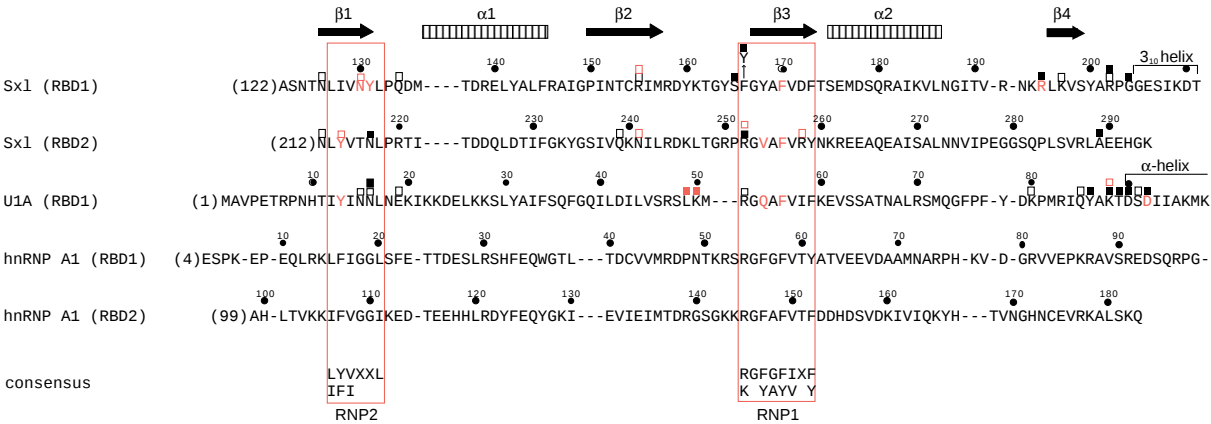


Figure 1 Sequence and secondary-structure alignment of the RBD(s) of Sxl, U1A and hnRNP A1 (refs 8–10). The RNP2 and RNP1 motifs, which correspond to the first and third β -strands, respectively, are indicated in red boxes, together with the consensus sequences, (L/I)-(Y/F)-(V/I)-X-X-L and (R/K)-G-(F/Y)-(G/A)-(F/Y)-(I/V)-X-(F/Y), respectively. The protein–RNA interactions of Sxl (this study) and U1A (ref. 12) are indicated as follows. Amino-acid residues involved in stacking

interactions with RNA are in red. Amino-acid residues that recognize the RNA backbone with their main chains and side chains are indicated by filled and open red squares, respectively. Amino-acid residues that recognize the bases with their main chains and side chains are indicated by filled and open black squares, respectively.

sequence specificities that are indistinguishable from those of the wild-type protein (data not shown). The crystals belong to the space group $I222$, with unit cell constants of $a = 77.9 \text{ \AA}$, $b = 86.8 \text{ \AA}$, $c = 160.4 \text{ \AA}$, and contain two complexes per asymmetric unit. The structure was solved by multiple isomorphous replacement augmented with anomalous dispersion (MIRAS), and was refined to $2.6 \text{ \AA}$ resolution with an R -factor of 20.1% ($R_{\text{free}} = 29.4\%$) (Table 1). Model building into the obtained electron density indicated that the first five uridines of the 17-mer RNA were degraded during crystallization, which is consistent with the band mobility of the RNA fragment from the complex crystals in the polyacrylamide-gel electrophoresis (PAGE) analysis (data not shown). We numbered the nucleotide residues of the RNA (GUUGUUUUUUU) from 1 to 12. The overall structure of the protein–RNA complex is shown in Fig. 2, and the correspondence between the secondary-structure elements and the amino-acid sequence of the protein is indicated in Fig. 1. In the electron density map, the side chains of three amino-acid residues in the loop between the $\beta 2$ and $\beta 3$ strands (the $\beta 2$ – $\beta 3$

loop) of RBD1 (Tyr 160, Lys 161 and Tyr 164), and one and five residues at the N- and C-termini, respectively, of the protein were not well ordered. As already mentioned, the asymmetric unit contains two copies of the complex. The first two nucleotide residues, G1 and U2, interact with another asymmetric complex, contributing to the crystal packing (data not shown). The structures of these asymmetric complexes are essentially identical, with a root-mean-square (r.m.s.) deviation of as little as $0.237 \text{ \AA}$ for all the 167 observed $\text{C}\alpha$ atoms. Nevertheless, the crystal-packing contacts occur at different sites between the two protein molecules in the asymmetric unit, so these contacts presumably had negligible effects on the structure of each molecule—for example, on the relative arrangement of the two RBDs. We now describe one of the two complexes in the asymmetric unit, whose overall structure is shown in Fig. 2. Mutagenesis and *in vitro* selection studies of target RNAs^{4,6,7} had not revealed how many and which nucleotide residues from the GUUGUUUUUUU sequence interact with Sxl: we find that the UGUUUUUUUU (U3–U11) region is continuously involved

Table 1 Crystallographic data and refinement statistics

Diffraction data	Native	Iodinated RNA	Seleno-Met	K ₂ PtCl ₄	K ₂ UO ₂ F ₅
Crystal					
Resolution ($\text{\AA}$)	2.6	2.7	2.7	2.7	2.8
Unique reflections	17,053	14,880	15,126	14,639	13,340
Total reflections	98,612	75,905	77,829	75,224	59,973
R_{merge} (%)†	8.9 (41.1)*	9.6	14.1	10.5	12.5
Completeness (%)	98.5 (87.4)*	97.4	98.7	95.9	97.5
Phasing statistics (50–3.0 $\text{\AA}$)					
R_{der} (%)‡		14.8	10.7	7.4	12.8
Overall phasing power§		1.68/1.37	0.68/0.60	0.32/0.21	0.30/0.22
R_{cullis}		0.71	0.92	0.98	0.97
Refinement statistics					
Resolution ($\text{\AA}$)	15 to 2.6				
Number of reflections	16,907				
Number of protein atoms	2,658				
Number of RNA atoms	494				
Number of water molecules	82				
R.m.s.d. bond lengths ($\text{\AA}$)	0.009				
R.m.s.d. bond angles ($^{\circ}$)	1.482				
R.m.s.d. impropers ($^{\circ}$)	0.649				
$R_{\text{work}}/R_{\text{free}}$ (%)¶	20.1/29.4				

The crystals belong to the space group $I222$, with $a = 77.9 \text{ \AA}$, $b = 86.8 \text{ \AA}$, $c = 160.4 \text{ \AA}$, and $\beta = 126^{\circ}$, and contain two complexes per asymmetric unit, resulting in a solvent content of 59% .
* Numbers in parentheses correspond to the values in the highest resolution shell.
† $R_{\text{merge}} = \sum_i \sum_h |\Sigma_j I_{ij} - \langle I_i \rangle| / \sum_i \sum_h I_{ij}$, where h are unique reflection indices, and i indicates symmetry equivalent indices.
‡ $R_{\text{der}} = \Sigma |F_{\text{PH}} - F_{\text{P}}| / \Sigma |F_{\text{PH}}|$, where $|F_{\text{PH}}|$ and $|F_{\text{P}}|$ refer to the measured structure factor amplitudes of the native and the derivative.
§ Phasing power is $f_{\text{rms}} / E_{\text{rms}}$, where $f_{\text{rms}} = [(\Sigma F_{\text{PH}}^2) / n]^{1/2}$ and $E_{\text{rms}} = [(\Sigma (F_{\text{PH}} - |F_{\text{P}} + f_{\text{H}}|)^2) / n]^{1/2}$.
|| $R_{\text{cullis}} = \Sigma (|F_{\text{H}}| - (|F_{\text{PH}}| - |F_{\text{P}}|)) / \Sigma (|F_{\text{PH}}| - |F_{\text{P}}|)$ (only for centric reflections), where $|F_{\text{H}}|$ represents the calculated heavy atom structure factor.
¶ $R_{\text{work}} = \Sigma |F_{\text{O}} - F_{\text{C}}| / \Sigma F_{\text{O}}$ for all reflections and $R_{\text{free}} = \Sigma |F_{\text{O}} - F_{\text{C}}| / \Sigma F_{\text{O}}$, calculated on the 10% of data excluded from refinement.

in the interaction with the protein (Fig. 2). In contrast, G1, U2 and U12 do not interact with the protein (Fig. 2). In the complex, the two RBDs sandwich the elongated RNA between their antiparallel β -sheet platforms (Fig. 2). This mode of interaction between the two-RBD protein and the single-stranded RNA contrasts with that between the single-RBD U1A fragment and the stem/loop-structured RNAs^{12,13}.

Domain arrangement of the two RBDs

The structures of the two RBDs in the RNA complex are essentially the same as those of the RNA-free solution structures of the isolated domains^{17,18}, except for some RNA-induced conformation changes

within the loop regions. The two RBDs interact with each other in the RNA complex. NMR of the RNA-free form of the Sxl RBD1-RBD2 indicates that the two RBDs do not interact with each other, and that the interdomain linker has no unique conformation (Y.M., unpublished results). In the complex, the two RBDs have their β -sheet platforms facing each other to form a V-shaped cleft, and the interdomain linker forms a short, distorted 3_{10} helix at the segment between Gly 205 and Thr 211 (Fig. 2). The eight β -strands of the two antiparallel β -sheets are spatially arranged, in the order RBD1 ($\beta 2$ – $\beta 3$ – $\beta 1$ – $\beta 4$)–RBD2($\beta 2$ – $\beta 3$ – $\beta 1$ – $\beta 4$) (Fig. 2). In the surface representation of the electrostatic potential, the interdomain cleft is strongly electropositive, which fits well with six nucleotide residues,

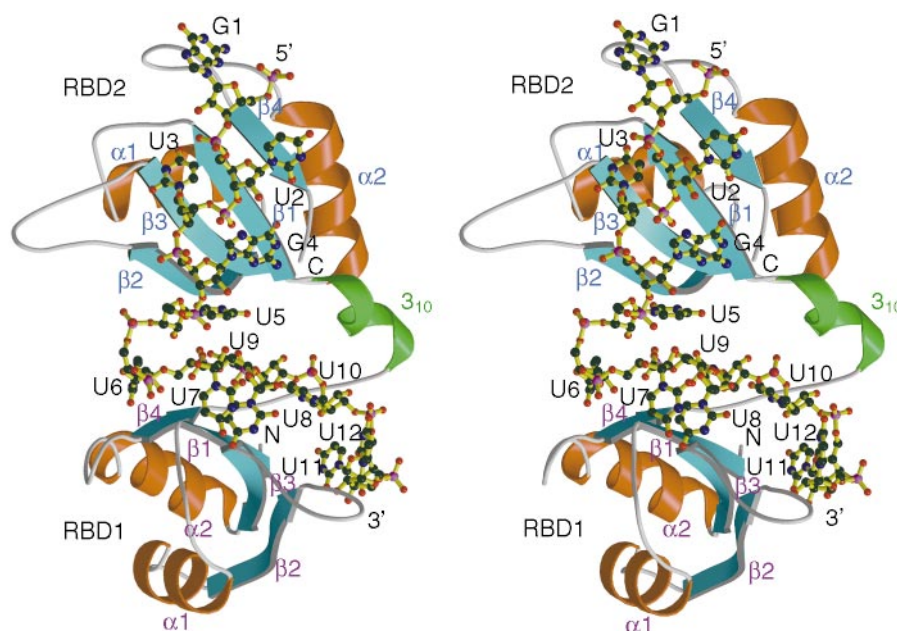


Figure 2 The structure of the complex between the Sxl protein (RBD1–RBD2) and the 12-nucleotide RNA derived from the *tra* polypyrimidine tract (stereo view). The overall folding of the protein is shown by a ribbon diagram. The α -helices are in red, the β -strands in cyan, the distorted 3_{10} helix in green, and the random coils in

grey. The secondary structure elements of RBD1, RBD2 and the interdomain linker are labelled in magenta, cyan and green, respectively. The bound RNA, which lacks secondary structure, is represented by a ball-and-stick model. This figure was drawn using MOLSCRIPT⁴⁸ and RASTER3D⁴⁹.

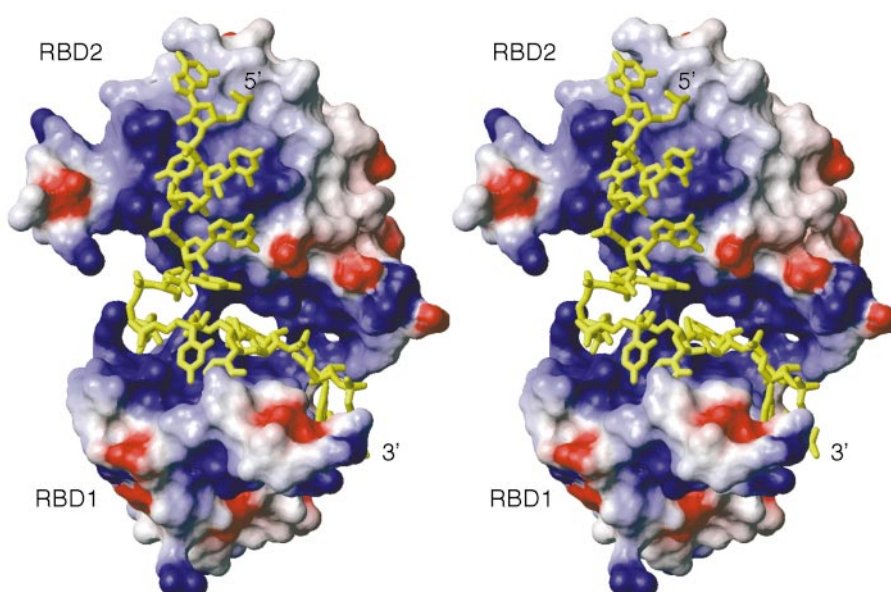


Figure 3 Distribution of the electrostatic potential on the solvent-accessible surface of Sxl (stereo view). The bound 12-nucleotide RNA is indicated in yellow by a ball-and-stick representation. The V-shaped interdomain cleft and the surface on the RBD2 β -sheet platform are strongly electropositive (blue); these

accommodate the phosphodiester backbone of the elongated RNA well. On the other hand, electronegative regions (red) are scattered. This figure was drawn using MOLMOL⁵⁰.

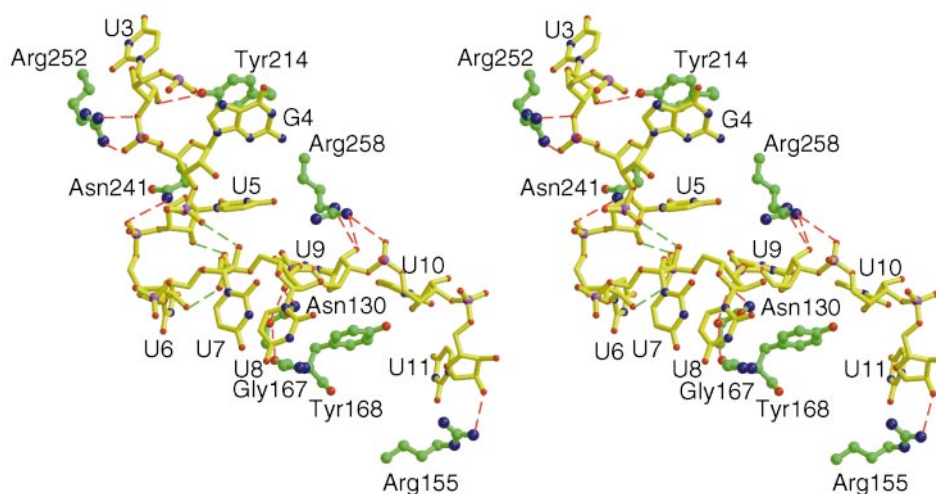


Figure 4 The U3–U11 (UGUUUUUUU) RNA segment in the complex (stereo view). RNA is represented by a wire-loop model in yellow, with atoms of nitrogen, oxygen and phosphorus shown in blue, red and purple, respectively. The ribose moieties, except for U8, are in the C2'-endo form. Light green broken lines show the

intramolecular hydrogen bonds that characterize the turn conformation of RNA. The side chains of the amino acids that interact with the RNA backbone are represented in green. Red broken lines denote the intermolecular hydrogen bonds of the amino-acid side chains with the phosphate-sugar backbone.

U6–U11, of the RNA backbone (Fig. 3). The other three nucleotide residues, U3–G4–U5, are bound to another positively charged surface on the RBD2 β -sheet platform (Fig. 3). Between RBD1 and RBD2, two interactions are observed, for N ζ of Lys 197 and the main-chain carbonyl of Val 238, and for O η of Tyr 131 and N ϵ of Gln 239 (not shown). Between RBD2 and the interdomain linker, two hydrogen bonds are formed: N ϵ H of Arg 202 and O δ of Asn 260, and the main-chain carbonyl of Ser 207 and the main-chain amide of Arg 262 (not shown). In contrast, there are no hydrogen bonds between RBD1 and the interdomain linker. Heterogeneous nuclear (hn)RNP A1 has two tandemly arranged RBDs, and the crystal structure of its two-RBD fragment (with no RNA) shows that the RNA-binding platforms are in the same plane, in an antiparallel orientation^{20,21}. This flat arrangement of the hnRNP A1 RBDs is completely different from the cleft-forming arrangement of the Sxl RBDs in a parallel orientation.

RNA conformation and backbone recognition

As shown in Figs 2 and 4, Sxl binding fixes the 9-nucleotide region of the RNA in a particular conformation that has no base pairs (neither G–U nor U–U). The U3–U11 segment is elongated in an irregular conformation, with a kink in the middle (Fig. 4). The torsion angles around U6 make a turn in the RNA backbone (Fig. 4), which is different from the turn structures found in other RNAs (refs 22–27). This turn around U6 is characterized by three direct internucleotide hydrogen bonds between the 2'-OH and phosphate groups: the 2'-OH groups of U5 and U6 provide hydrogen bonds to the phosphate group of U8; and the 2'-OH group of U7 hydrogen-bonds to the phosphate group of U5 (Fig. 4). The turn conformation is also characterized by the only intramolecular base-stacking interaction, between U7 and U8, whereas none of the other RNA bases participates in intramolecular stacking. Moreover, it is remarkable that the sugar conformations of all of the nucleotide residues, except for U8, are the C2'-endo form. Correspondingly, the phosphate–phosphate distance at U8 is appreciably shorter than those at other nucleotide residues (data not shown). These sugar puckerings in the crystal are consistent with those determined for the same protein–RNA complex in solution by isotope-aided NMR spectroscopy (I.K. and Y.M., unpublished results). This unusually C2'-endo-rich RNA conformation enables the bases to be exposed to the protein. In contrast, in the structure of the hairpin RNA in the U1A complex, only a few sugars are in the C2'-endo form, and all but one of the bases participate in the intramolecular stacking

interactions¹².

The phosphate–sugar backbone of the UGUUUUUUU (U3–U11) region interacts extensively with the β -sheet platforms of the two RBDs (Fig. 2), in contrast to the U1A–RNA complexes, which have only a few interactions between the protein and the RNA backbone of the single-stranded region^{12,13}. A long length of the RNA backbone interacts with the amino-acid side chains of RBD2 (Fig. 4) through five direct hydrogen bonds and two salt bridges: the hydrogen bonds of the O2' and O3' groups of U3 with Tyr 214 and Arg 252, respectively (Fig. 5a), the salt bridge of the phosphate-group oxygen of G4 with Arg 252 (Fig. 5a), the hydrogen bond of the OP of U6 with Asn 241 (Fig. 5c), the hydrogen bond of the O2' of U9 with Arg 258, and the salt bridge of the OP of U10 with Arg 258 (Fig. 5e). Congruent with these findings, the replacement of Arg 252 with alanine has been reported to abolish the RNA-binding activity of Sxl (ref. 28). On the other hand, only the 3'-end region of the RNA interacts with RBD1, by two direct hydrogen bonds and two water-mediated hydrogen bonds (Fig. 4): the phosphate-group oxygen of U9 forms a hydrogen bond with the Asn 130 side chain and two water-mediated hydrogen bonds with the main-chain amino and carbonyl groups of Tyr 131 (Fig. 5e) and Gly 167 (Fig. 4), respectively; the O2' of U11 forms a hydrogen bond with the Arg 155 side chain (Fig. 5f). In contrast, the sugar–phosphate moieties of U5, U7 and U8, which are involved in a characteristic turn conformation with the internucleotide hydrogen bonds, do not interact with the protein. Thus, six of the nine 2'-OH groups are involved in inter- and/or intramolecular interactions, which may explain the 10⁴-fold weaker binding of a DNA fragment of the same nucleotide sequence with a Sxl RBD1–RBD2 fragment¹⁵.

In contrast, in the U1A–RNA complexes, the RNA backbone interacts only locally with the protein^{12,13}. This is probably because the RNA conformation primarily depends on the base-paired double-stranded region(s) of the hairpin or internal-loop secondary structure, which is required for RNA binding by the U1A protein²⁹. Arg 52 of U1A recognizes the guanine base of the loop-closing G16–C5 pair (Fig. 6), whereas the corresponding Arg 252 of the Sxl RBD2 interacts extensively with the RNA backbone (Figs 4, 5a, 6).

RNA-base recognition

We found that Sxl RBD1 and RBD2 recognize the bases of the UGUUUUUUU (U3–U11) region. All of the nine bases are bound to their distinct recognition sites of the protein; the N3H and/or O4 groups of the eight uracil moieties and the 2-amino group of

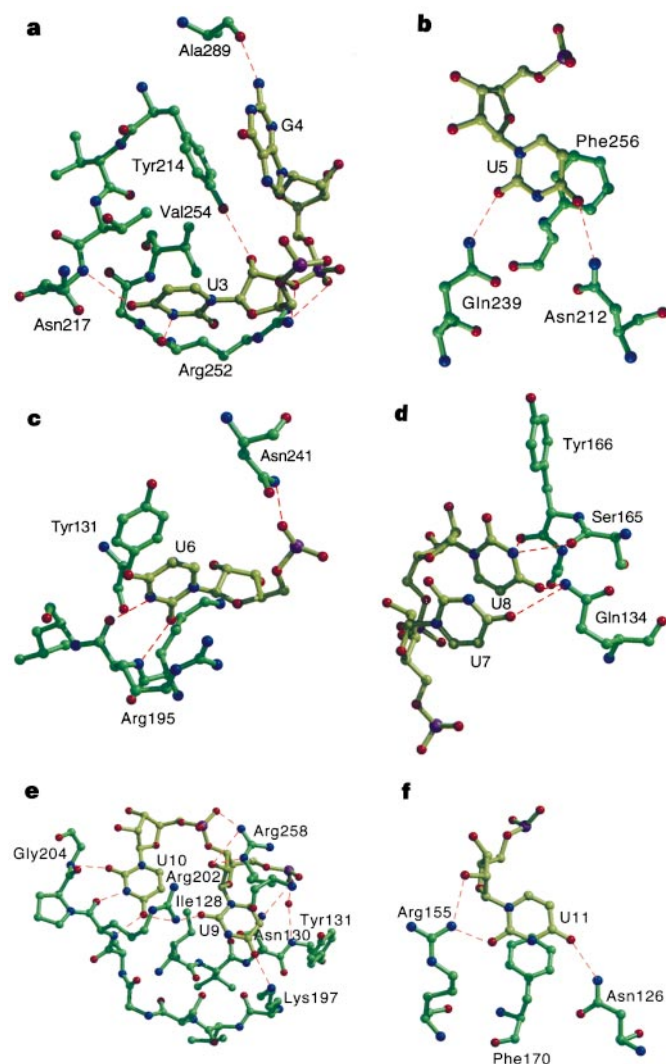


Figure 5 Recognition of U3–U11 (UGUUUUUUU) by the Sxl RBD1 and RBD2 domains. The carbon atoms of the protein and the RNA are shown in dark and light green, respectively. Hydrogen-bond interactions between RNA and the protein are indicated by broken red lines. **a**, Recognition of U3 and G4 on RBD2; **b**, U5 recognition on RBD2; **c**, U6 recognition on RBD1; **d**, Recognition of U7 and U8 that are stacked on each other on RBD1; **e**, Recognition of U9 and U10 on RBD1; **f**, U11 recognition on RBD1.

the guanine moiety are recognized (see later). The UGU region is recognized by RBD2, whereas the following UUUUUU region is recognized mainly by RBD1. It should be noted here that the UGUUUUUUU sequence is recognized continuously. Specifically, no nucleotide residue can be inserted between the last RBD2-bound residue (U5) and the first RBD1-bound residue (U6), because the characteristic turn structure of the RNA conformationally hinders a one-nucleotide insertion (Fig. 4), and spatially prevents the insertion of two or more nucleotides (Fig. 3). Therefore, the RNA recognition by Sxl is strictly specific to the UGUUUUUUU sequence. These results establish the mechanisms by which uridine-rich sequences with one or more cytidine residues bind only weakly to the Sxl protein^{3,4,6}. In contrast, an essential splicing factor, U2AF, has three RNP domains, and prefers cytidine-containing, uridine-rich polypyrimidine tracts to that of the *tra* pre-mRNA polypyrimidine tract, which contains the UGUUUUUUU region⁷. Thus, Sxl binding prevents the *tra* pre-mRNA polypyrimidine tract from binding U2AF, and accordingly, the distal, female-specific splice site is activated by U2AF⁵. It would be interesting to compare the polypyrimidine-tract recognition mechanism of the general

	Sxl RBD1			Sxl RBD2			U1A		
	Asn 126	H	U11	Asn 212	H	U5	Thr 11	H	Ser91
	Ile 128	S	U9, U10	Tyr 214	S	G4, U3	Tyr 13	S	C10
β1	Asn 130	H	U9	Thr 216	-	-	Tyr 13	H	Gln54
	Asn 130	S	U9	-	-	-	Asn 15	H	G9
	Tyr 131	S	U6	Asn 217	H	U3	Asn 16	H	U8, G9
β1-α1	Gln 134	H	U7, U8	Arg 220	-	-	Asn 16	H	U7, G9
	Thr 153	-	-	Gln 239	H	U5	Asp 42	-	-
β2	Arg 155	H	U11, U11	Asn 241	H	U6	Leu 44	-	-
	Tyr 160	-	-	Lys 246	-	-	Leu 49	H	G16
	Lys 161	-	-	Leu 247	-	-	Lys 50	H	G9
β2-β3	Ser 165	H	U8	Pro 251	-	-	-	-	-
	Tyr 166	H	U8	Arg 252	H	U3	Arg 52	H	A6, G16
β3	Tyr 168	S	U9, U10	Val 254	S	U3	Gln 54	S	G9
	Phe 170	S	U11	Phe 256	S	U5	Phe 56	S	A11
	Asp 172	-	-	Arg 258	H	U9, U10	Ile 58	S	Ile93, Ile94
α2-β4	Lys 194	-	-	Gln 282	-	-	Lys 80	H	U8
	Arg 195	S	U6	Pro 283	-	-	Pro 81	-	-
β4	Arg 195	H	U6	Ser 285	-	-	Arg 83	H(w)	U8
	Lys 197	H	U9	Arg 287	-	-	Gln 85	H	C10
	Ser 199	-	-	Leu 288	-	-	Tyr 86	H	C10
	Tyr 200	-	-	Ala 289	H	G4	Lys 87	-	-
	Ala 201	-	-	Glu 290	-	-	Lys 88	H	C10
	Arg 202	H	U9, U10	-	-	-	Lys 88	H	C10
	Arg 202	H	U10	Glu 291	-	-	Thr 89	H	A11
	Pro 203	-	-	His 292	-	-	Asp 90	H	C12
	Gly 204	H	U10	Gly 293	-	-	Ser 91	H	A11
	Gly 205	-	-	Lys 294	-	-	Asp 92	S	C12
	Glu 206	-	-	-	-	-	Asp 92	H	C12

Figure 6 Protein–RNA interactions of the Sxl and U1A RBDs. The protein–RNA interactions in the Sxl RBD1–RBD2–RNA complex (this study) and the U1A RBD–RNA complex (ref. 12) are summarized along the secondary structure: the β1 strand (RNP2), the β1–α1 loop, the β2 strand, the β2–β3 loop, the β3 strand (RNP1), the α2–β4 loop, the β4 strand and the C-terminal region of RBD (indicated on the left). Each amino-acid residue of one RBD is listed in the same row as those at the corresponding secondary-structural positions of the other two RBDs. In the three columns shown for each RBD, the type of interaction is indicated as H for hydrogen bond or salt bridge, or as S for stacking, hydrophobic or van der Waals interaction (a dash signifies that there is no interaction); the nucleotide residue(s) in the third column interact(s) with the amino-acid residue in the first column. The main chains and side chains of amino-acid residues are shown in black, italic or blue letters, respectively; the phosphate–ribose moiety and the base of a nucleotide residue are distinguished similarly.

splicing factor U2AF with that of the Sxl protein, a strictly sequence-specific splicing repressor.

U3–G4–U5 recognition. The bases of the UGU (U3–U5) region are located on the RBD2 β-sheet surface (Fig. 2). The uracil base of U3 contacts the side chain of Val 254 (the third residue in the RNP1 motif; Fig. 1), whereas the G4 and U5 bases stack on the aromatic rings of Tyr 214 (the second residue in RNP2) and Phe 256 (the fifth residue in RNP1), respectively (Fig. 5a, b). In the case of the Sxl RBD2, the N3H and O4 groups of U3 form hydrogen bonds to the main-chain carbonyl group of Arg 252 in RNP1 and the main-chain amide group of Asn 217 in RNP2, respectively (Fig. 5a). The 2-amino group of G4 forms a hydrogen bond to the main-chain carbonyl of Ala 289 (Fig. 5a). The O2 and O4 groups of U5 provide hydrogen bonds to the side chains of Gln 239 and Asn 212, respectively (Fig. 5b). The locations of the U3, G4, and U5 bases on the Sxl RBD2 β-sheet are similar to those of the G9, C10, and A11 bases, respectively, of the hairpin RNA on the U1A β-sheet (not shown). These similarities are probably due to the similar sequences of the RNP2 and RNP1 motifs (except for the third residue of RNP1) (Fig. 1)¹⁸. However, the specific RNA-recognition mechanisms are different between the Sxl RBD2 and the U1A RBD; the correspond-

ing amino-acid residues play different roles at many positions (Fig. 6). As a good example, the Arg252 side chain of the Sxl RBD2 forms hydrogen bonds with the backbone moieties of U3 and G4, whereas the corresponding Arg 52 side chain of the U1A RBD recognizes the bases of A6 and G16 (Fig. 6).

UUUUUU (U6–U11) recognition. This U₆ region is recognized mainly by the Sxl RBD1. One side of U6 base stacks on the aromatic ring of Tyr 131 (the fifth residue in RNP2), while its other side base stacks with the guanidino group of Arg 195 in the α 2- β 4 loop of RBD1 (Fig. 5c). The O2 and N3H groups of U6 form hydrogen bonds to the main-chain amide and carbonyl groups, respectively, of Arg 195 (Fig. 5c). The U7 and U8 bases are stacked on each other, and the two O4 groups of U7 and U8 form hydrogen bonds to the side chain of Gln 134 (Fig. 5d). In addition, the N3H group of U8 forms hydrogen bonds to the main-chain carbonyl groups of Ser 165 and Tyr 166 (Fig. 5d). The U9 base is specifically recognized by RBD1 and the interdomain linker. First, the U9 base contacts the side chain of Asn 130 (the fourth residue in RNP2) (Fig. 5e). The O2 and O4 groups of U9 provide hydrogen bonds to the side chains of Arg 202 and Lys 197, respectively (Fig. 5e). The U10 base is closely recognized by the interdomain linker; the O2 and N3H groups of U10 form hydrogen bonds to the main-chain amide group of Gly 204 and the main-chain carbonyl group of Arg 202, respectively; and the O4 group of U10 forms hydrogen bonds to the main-chain amide group and the side chain of Arg 202 (Fig. 5e). In addition, the side chain of Tyr 168 (the third residue in RNP1) hydrophobically interacts with the ribose moieties of U9 and U10 (Fig. 4). Moreover, the Sxl RBD2 cooperates with the Sxl RBD1 in the binding of U9 and U10; the side chain of Arg 258 (RNP1) of the Sxl RBD2 forms three hydrogen bonds to the sugar–phosphate backbone of U9–U10, as already described (Figs 4, 5e). Thus, on the Sxl RBD1, the U9 and U10 bases properly contact the small aliphatic side chain of Ile 128 (the second residue in RNP2) (Fig. 5e). The second and fifth amino-acid residues of RNP2 of the Sxl RBD1 are characteristically different from those of the Sxl RBD2 and the U1A RBD (Fig. 1)¹⁸. Therefore, none of the interactions of Sxl RBD1 amino-acid residues with the U6–U10 region are conserved in the RNA interactions of the Sxl RBD2 and the U1A RBD (Fig. 6). In contrast to the distinct interactions of U6–U10 with RBD1, the interaction of U11 with RBD1 (Fig. 5f) is similar to that of U5 with RBD2 (Fig. 5b). The uracil base of U11 stacks on the aromatic ring of Phe 170 (the fifth residue in RNP1) (Fig. 5f). The O2 and O4 groups of U11 form hydrogen bonds to the side chains of Arg 155 and Asn 126, respectively (Fig. 5f), while those of U5 form hydrogen bonds to the side chains of Gln 239 and Asn 212, respectively (Fig. 5b).

ELAV-family proteins. The ELAV-family proteins have two RBDs whose amino-acid sequences are both highly homologous to those of the Sxl RBD1 and RBD2. The characteristic RNP2 sequence of RBD1 and other RNA-interacting amino-acid residues of the Sxl protein identified here are well conserved in the ELAV family. The RNA-binding modes of these ELAV-family proteins are therefore likely to be similar to the RNA-recognition mechanism of the Sxl protein, although their nucleotide-sequence specificities are not exactly the same. The *Drosophila* ELAV protein is involved in regulating alternative splicing that is essential for neuronal development³⁰. In the cases of the vertebrate ELAV-family proteins, such as human Hel-N1, HuC, HuD and HuR, their first two RBDs bind specifically to the A+U-rich elements (AREs) located in the 3' UTRs of mRNAs that encode growth-regulating proteins, including c-Fos and c-Myc, and the mRNAs are therefore stabilized by a reduction in the ARE-directed rapid mRNA decay^{31,32}. The *Xenopus* ElrA protein binds to an embryonic cytoplasmic polyadenylation element, which consist of at least 12 uridines³⁰.

Base-paired or single-stranded RNA

The RNA recognition mechanism of Sxl is different from those of the U1A and U2B' snRNP proteins^{11–13}. These snRNP RBDs depend

fundamentally on the higher-order structure or the intramolecular base pairs of the target RNA, with the recognition sequence being in a single-stranded loop region; the U1A protein, for example, has a much weaker (10^4 -fold) affinity for a single-stranded 23-mer RNA with no base pairs, even though the proper recognition sequence is present²⁹. The double-stranded stem has two important roles: (1) it restricts the conformation of the RNA loop and reduces the entropy loss accompanying protein binding; and (2) the C-G base pair that closes the RNA loop is crucial for positioning the RNA¹². The requirement of the double-stranded region of RNA for base-sequence-specific protein binding has been a general finding for the limited number of RNA–protein complex structures determined so far by X-ray crystallography and/or NMR spectroscopy; all of the RNAs have secondary structure(s) involving base-paired stem region(s). In the crystal structures of aminoacyl-tRNA synthetases complexed with their cognate tRNAs, the synthetase binds to one or more stem regions of the L-shaped tertiary structure of the tRNA, and recognizes some base pairs in the stem regions and/or unpaired bases in the single-stranded regions, such as the anticodon^{33–37}. Structures of viral proteins complexed with their target RNAs have revealed that the β -sheet surface interacts with the backbone of the stem and the exposed bases in a stem–loop hairpin RNA structure²³, and that a peptide segment binds in the major groove of the base-paired stem^{24–27}.

However, there are many proteins that can recognize a single-stranded region of a target RNA, without the assistance of the base-paired stems. Many RBDs recognize such single-stranded RNAs³⁸. Certain single-stranded DNAs are also sequence-specifically recognized by proteins containing RBDs^{39–41}. Our RNA–protein complex structure provides a structural basis for the specific recognition of a long and continuous sequence of RNA (and/or DNA) with no base pairs.

Methods

Crystallization and data collection. Wild-type and mutant RBD1–RBD2 (122–294) portions of the Sxl protein were expressed in *Escherichia coli* and purified as described⁴². Oligoribonucleotides were chemically synthesized and purified as described⁴². The best crystals were grown at 20 °C by the hanging-drop vapour-diffusion method (protein at 4–10 mg ml^{−1}; RNA-to-protein ratio, 1–1.2:1) against a reservoir solution containing 10% PEG6000, 50 mM NaMOPS (pH 7.0), 100 mM Li₂SO₄, 3 mM spermine, 1–3% dioxane, 2 mM Zn(CH₃COO)₂ and 2 mM MgCl₂. A selenomethionine (SeMet)-substituted complex, a complex with (5)I-U6-containing RNA, and platinum and uranium derivatives were used to solve the phase problem by the MIRAS method (Table 1). The SeMet-substituted protein was prepared by growing the *E. coli* methionine-auxotrophic strain B834, transformed with overexpression vector, in minimal medium, in which methionine was replaced by SeMet. The 17-nucleotide RNA, carrying a substitution of U6 by 5-iodo-uridine, was synthesized and purified as before⁴². Crystals of the SeMet-substituted complex were grown by cross-seeding. Platinum and uranium derivatives were prepared by soaking crystals in K₂PtCl₄ and in K₃UO₂F₅. Data were collected at 20 °C on a Raxis-IV image plate detector mounted on a Rigaku X-ray source.

Phasing and refinement. All data were processed using the DENZO⁴³ and SCALEPACK⁴³ programs with a 1 σ cutoff. Subsequent phase calculations were carried out with the CCP4 program suite⁴⁴. Two iodine sites were determined by the RSPS⁴⁴ program from the isomorphous and anomalous difference Patterson maps. The initial phases from the iodine derivative were used to locate the positions of the platinum, uranium and selenium atoms by difference Fourier analysis. Heavy-atom parameters were refined using the MLPHARE⁴⁴ program: the resulting MIRAS map was of excellent quality and showed a clear solvent boundary. The electron density map was further improved by density-modification procedures such as solvent flattening, histogram matching, solvent flipping, and NCS averaging, using program DM⁴⁴. Model building in the electron density map was done using program O (ref. 45). The position of the iodine guided the assignment of the RNA residues in the electron density map. Crystallographic positional and slow-cooling refinements were carried out against the 2.6 Å data set using X-PLOR⁴⁶ (Table 1). The final models have

very good geometry, as examined by PROCHECK⁴⁷: 89.7% of the residues have ϕ/ψ angles in the 'most favoured region' of the Ramachandran plot and 99.0% are in the 'allowed regions'.

Received 20 November 1998; accepted 10 March 1999.

- Harford, J. B. & Morris, D. R. *mRNA Metabolism and Post-transcriptional Gene Regulation* (Wiley-Liss, New York, 1997).
- Cline, T. W. & Meyer, B. J. Vive la difference: males vs females in flies vs worms. *Annu. Rev. Genet.* **30**, 637–702 (1996).
- Sosnowski, B. A., Belote, J. M. & McKeown, M. Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell* **58**, 449–459 (1989).
- Inoue, K., Hoshijima, K., Sakamoto, H. & Shimura, Y. Binding of the *Drosophila* Sex-lethal gene product to the alternative splice set of *transformer* primary transcript. *Nature* **344**, 461–463 (1990).
- Valcárcel, J., Singh, R., Zamore, P. D. & Green, M. R. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature* **362**, 171–175 (1993).
- Sakashita, E. & Sakamoto, H. Characterization of RNA binding specificity of the *Drosophila* sex-lethal protein by *in vitro* ligand selection. *Nucleic Acids Res.* **22**, 4082–4086 (1994).
- Singh, R., Valcárcel, J. & Green, M. R. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* **268**, 1173–1176 (1995).
- Burd, C. G. & Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615–621 (1994).
- Birney, E., Kumar, S. & Krainer, A. R. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing. *Nucleic Acids Res.* **21**, 5803–5816 (1993).
- Varani, G. & Nagai, K. RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 407–445 (1998).
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**, 515–520 (1990).
- Oubridge, C., Ito, N., Evans, P., Teo, C.-H. & Nagai, K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438 (1994).
- Allain, F. H.-T. *et al.* Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. *Nature* **380**, 646–650 (1996).
- Price, S. R., Evans, P. R. & Nagai, K. Crystal structure of the spliceosomal U2B'-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **394**, 645–650 (1998).
- Kanaar, R., Lee, A. L., Rudner, D. Z., Wemmer, D. E. & Rio, D. C. Interaction of the Sex-lethal RNA binding domains with RNA. *EMBO J.* **14**, 4530–4539 (1995).
- Sakashita, E. & Sakamoto, H. Protein-RNA and protein-protein interactions of the *Drosophila* Sex-lethal mediated by its RNA-binding domains. *J. Biochem.* **120**, 1028–1033 (1996).
- Samuels, M., Deshpande, G. & Schedl, P. Activities of the Sex-lethal protein in RNA binding and protein-protein interactions. *Nucleic Acids Res.* **26**, 2625–2637 (1998).
- Inoue, M. *et al.* A characteristic arrangement of aromatic amino acid residues in the solution structure of the amino-terminal RNA-binding domain of *Drosophila* Sex-lethal. *J. Mol. Biol.* **272**, 82–94 (1997).
- Lee, A. L., Kanaar, R., Rio, D. C. & Wemmer, D. E. Resonance assignments and solution structure of the second RNA-binding domain of Sex-lethal determined by multidimensional heteronuclear magnetic resonance. *Biochemistry* **33**, 13775–13786 (1994).
- Shamoo, Y., Krueger, U., Rice, L. M., Williams, K. R. & Steitz, T. A. Crystal structure of the two RNA-binding domains of human hnRNP A1 at 1.75 Å resolution. *Nature Struct. Biol.* **4**, 215–222 (1997).
- Xu, R.-M., Jokhan, L., Cheng, X., Mayeda, A. & Krainer, A. R. Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **5**, 559–570 (1997).
- Simons, R. W. & Grunberg-Manago, M. *RNA Structure and Function*. (Cold Spring Harbor Laboratory Press, New York, 1998).
- Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* **371**, 623–626 (1994).
- Puglisi, J. D., Chen, L., Blanchard, S. & Frankel, A. D. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **270**, 1200–1203 (1995).
- Battiste, J. L. *et al.* α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273**, 1547–1551 (1996).
- De Guzman, R. N. *et al.* Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ -RNA recognition element. *Science* **279**, 384–388 (1998).
- Cai, Z. *et al.* Solution structure of P22 transcriptional antitermination N peptide-box B RNA complex. *Nature Struct. Biol.* **5**, 203–212 (1998).
- Lee, A. L. *et al.* Chemical shift mapping of the RNA-binding interface of the multiple-RBD protein Sex-lethal. *Biochemistry* **36**, 14306–14317 (1997).
- Hall, K. B. Interaction of RNA hairpins with the human U1A N-terminal RNA binding domain. *Biochemistry* **33**, 10076–10088 (1994).
- Good, P. J. The role of *elav*-like genes, a conserved family encoding RNA-binding proteins, in growth and development. *Semin. Cell Dev. Biol.* **8**, 577–584 (1997).
- Fan, X. C. & Steitz, J. A. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the *in vivo* stability of ARE-containing mRNAs. *EMBO J.* **17**, 3448–3460 (1998).
- Peng, S. S.-Y., Chen, C.-Y. A., Xu, N. & Shyu, A.-B. RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.* **17**, 3461–3470 (1998).
- Rould, M. A., Perona, J. J. & Steitz, T. A. Structural basis of anticodon loop recognition by glutamyl-tRNA synthetase. *Nature* **352**, 213–218 (1991).
- Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. & Moras, D. Yeast tRNA^{Asp} recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature* **362**, 181–184 (1993).
- Biou, V., Yaremchuk, A., Tukalo, M. & Cusack, S. The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA^{Ser}. *Science* **263**, 1404–1410 (1994).
- Cusack, S., Yaremchuk, A. & Tukalo, M. The crystal structures of *T. thermophilus* lysyl-tRNA synthetase complexed with *E. coli* tRNA^{Lys} and a *T. thermophilus* tRNA^{Lys} transcript: anticodon recognition and conformational changes upon binding of a lysyl-adenylate analogue. *EMBO J.* **15**, 6321–6334 (1996).
- Cusack, S., Yaremchuk, A., Krikiliviy, I. & Tukalo, M. tRNA^{Phe} anticodon recognition by *Thermus thermophilus* prolyl-tRNA synthetase. *Structure* **6**, 101–108 (1998).
- Shamoo, Y., Abdul-Manan, N. & Williams, K. R. Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res.* **23**, 725–728 (1995).
- McKay, S. J. & Cooke, H. hn RNP A2/B1 binds specifically to single stranded vertebrate telomeric repeat TTAGGG_n. *Nucleic Acids Res.* **20**, 6461–6464 (1992).
- Kajita, Y., Nakayama, J., Aizawa, M. & Ishikawa, F. The UUAG-specific RNA binding protein, heterogenous nuclear ribonucleoprotein D0. *J. Biol. Chem.* **270**, 22167–22175 (1995).
- Sarig, G., Weisman-Shomer, P., Erlitzki, R. & Fry, M. Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes. *J. Biol. Chem.* **272**, 4474–4482 (1997).
- Kim, I. *et al.* NMR analysis of the hydrogen bonding interactions of the RNA-binding domains of the *Drosophila* Sex-lethal protein with target RNA fragments with site-specific [¹⁵N]uridine substitutions. *Nucleic Acids Res.* **25**, 1565–1569 (1997).
- Otwinski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- CCP4: Collaborative Computational Project No. 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Jones, T. A., Zhou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta. Crystallogr. A* **47**, 110–119 (1991).
- Brünger, A. T. *X-PLOR: A System for X-Ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
- Kraulis, P. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Merritt, E. A. & Bacon, D. J. Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).
- Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55, 29–32 (1996).

Acknowledgements. We thank D. G. Vasylyev for helpful discussions, and M. Horikoshi for data collection. This work was supported in part by Grants-in-Aid for Scientific Research on Priority Areas to S.Y. from the Ministry of Education, Science, Culture and Sports of Japan.

Correspondence and requests for materials should be addressed to S.Y. (e-mail: yokoyama@y-sun.biochem.s.u-tokyo.ac.jp). The coordinates for the Sxl-RNA complex have been deposited in the Brookhaven Protein Data Bank, accession code 1B7F.

Spectroscopic identification of a galaxy at a probable redshift of $z = 6.68$

Hsiao-Wen Chen, Kenneth M. Lanzetta & Sebastian Pascarelle

Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York 11794-3800, USA

The detection and identification of distant galaxies is an important goal of observational cosmology, as such galaxies are seen at a time when the Universe was very young. The development of new techniques and instrumentation permits the search for ever-fainter galaxies, and so aids attempts to determine when the first stars and galaxies formed. Here we report the identification of a galaxy at a probable redshift of 6.68, the most distant object yet detected. The galaxy's spectrum is characterized by an abrupt discontinuity at a wavelength $\lambda \approx 9,300 \text{ \AA}$, which we interpret as arising from the absorption of light at shorter wavelengths by hydrogen gas along the line of sight (the Lyman- α decrement), and by an emission line at $\lambda \approx 9,334 \text{ \AA}$, which we interpret as the Lyman- α line at a redshift of 6.68. The galaxy is relatively bright: the ultraviolet luminosity density contributed by this one galaxy is almost ten times the value measured at $z = 3$.

At near-infrared wavelengths, where background sky light is the dominant source of noise, the Hubble Space Telescope (HST) using the Space Telescope Imaging Spectrograph (STIS) is more sensitive than the Keck telescope. This is because, first, the sky is considerably fainter from space, and second, considerably higher spatial resolution is attained from space, thus admitting far less contaminating sky light. To exploit this sensitivity of STIS at near-infrared wavelengths, the Space Telescope Science Institute and the STIS instrument team at the Goddard Space Flight Center initiated the STIS Parallel Survey, in which deep STIS observations are obtained in parallel with other observations¹. We selected for analysis very deep observations acquired in slitless spectroscopy mode (which records the dispersed light of an entire field of view), because these observations are best suited for identifying distant galaxies.

The very deep observations were obtained by HST using STIS from 23 to 26 December 1997 towards a region of sky flanking the Hubble Deep Field. Most observations consisted of a pair of images: a direct image taken using no filter and a dispersed image taken using the G750L grating. Additional observations consisted of only a direct image. The integration time of the direct images totalled 4.5 h over 82 exposures, and the integration time of the dispersed images totalled 13.5 h over 60 exposures.

We summed the direct and dispersed images using conventional image-processing techniques. First, we reduced the images using standard software and applied corrections for flat-field variations and illumination pattern. Next, we registered each pair of images to a common origin, using pointing offsets determined from positions of bright stars in the direct images. We then measured the noise characteristics of the images, which are set by the background sky level and the readout noise. (The background sky level varies by as much as a factor of two between individual exposures.) Finally, we summed the direct images to form a summed direct image and summed the dispersed images to form a summed dispersed image, using optimal weights determined from the measured noise characteristics and rejecting deviant pixel values caused by cosmic-ray events and bad pixels. The spatial resolution (full-width at half-maximum) of the summed direct and dispersed images is $\sim 0.08 \text{ arcsec}$, the 1σ single pixel detection threshold of the summed direct image is $\sim 26.2 \text{ mag arcsec}^{-2}$, and the 1σ single

pixel detection threshold of the summed dispersed image at $\lambda \approx 9,800 \text{ \AA}$ is $\sim 5.5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ arcsec}^{-1}$.

We extracted one-dimensional spectra from the summed dispersed image using a new spectrum extraction technique. The extraction is especially difficult because the image of the field is covered by the light of faint galaxies, which, when dispersed, overlaps in a complicated way. To overcome this difficulty, we used the summed direct image to determine not only the exact locations but also the exact two-dimensional spatial profiles of the spectra on the summed dispersed image. These spatial profiles are crucial because (1) they provide the 'weights' needed to optimally extract the spectra, and (2) they provide the models needed to deblend the overlapping spectra and determine the background sky

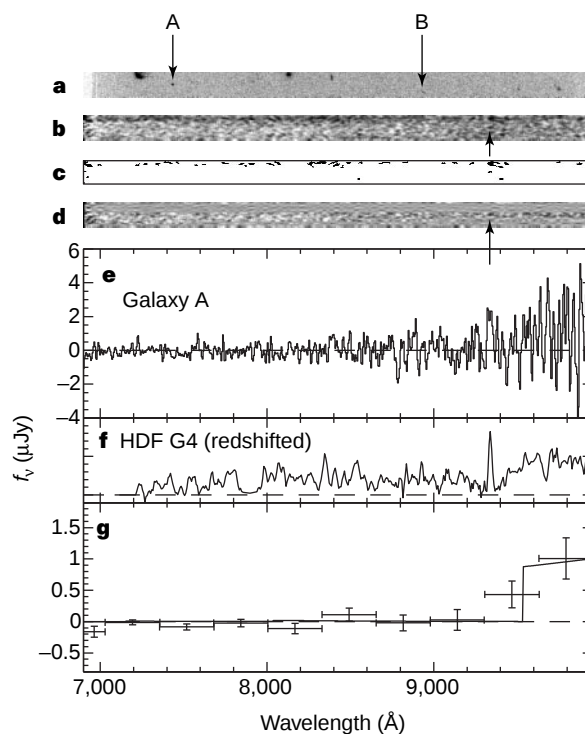


Figure 1 Image and spectra of galaxy A and its neighbours. **a**, The direct image of galaxy A and its neighbours, with arrows pointing to galaxies A and B. Angular extent of the image is 51 arcsec wide by 2 arcsec high. **b**, The dispersed image of galaxy A and its neighbours, with an arrow pointing to Ly α emission line of galaxy A. Pixel size is $4.882 \text{ \AA}$ in the horizontal direction and 0.1 arcsec in the vertical direction, resolution is roughly two pixels (in both directions), and spectrum is boxcar-smoothed by five pixels in the spectral direction and two pixels in the spatial direction. **c**, Map of the detected emission lines. (We note that the continua of the topmost bright objects have been broken up by the SExtractor program and are not individual emission lines.) **d**, The dispersed image with best-fit model spatial profiles of all objects but galaxy A subtracted, with an arrow pointing to the Ly α emission line. **e**, the one-dimensional spectrum of galaxy A. Pixel size is $5 \text{ \AA}$, resolution is roughly two pixels, and spectrum is boxcar-smoothed by three pixels. **f**, Redshifted (to $z = 6.68$) spectrum of a moderate-redshift ($z = 4.421$) galaxy. **g**, the one-dimensional spectrum of galaxy A cast into $325\text{-\AA}$ bins together with best-fit spectrophotometric template spectrum. Vertical error bars indicate 1σ uncertainties and horizontal error bars indicate bin sizes. The images were recorded by a $1,024 \times 1,024$ CCD detector, which for the direct images was operated in unbinned mode (resulting in a $1,024 \times 1,024$ grid) and for the dispersed images was operated in 1×2 binned mode (resulting in a $1,024 \times 512$ grid). The spatial scale of the image on the detector was $0.05 \text{ arcsec pixel}^{-1}$, which yielded a field of view of $51 \times 51 \text{ arcsec}^2$. The telescope was displaced by several arcsec between each observation in order to reduce the effects of pixel-to-pixel sensitivity variations.

level. First, we identified objects in the summed direct image, using the SExtractor program². Roughly 250 objects were identified in the summed direct image. Next, we modelled each pixel of the summed dispersed image as a linear sum of contributions from relevant portions of all overlapping neighbouring objects and from background sky. The model parameters included $\sim 250,000$ 'object' parameters (from roughly 1,000 spectral pixels each of about 250 objects) and 5,000 'sky' parameters (from approximately 1,000 fourth-order polynomials). Finally, we minimized χ^2 between the model and the data with respect to the model parameters to form one-dimensional spectra. (In practice, we imposed a condition of smoothness on the spectral variation of the sky in order to reduce the number of sky parameters. We first smoothed the model of the sky by 10 pixels in the spectral direction, and we subtracted this model from the data. Then we performed the extraction again, but this time with sky subtraction turned off. In this way we reduced the effective number of sky parameters from $\sim 5,000$ to more like 500. We found by experimentation that this number of parameters is required to represent the spatial and spectral variations of the sky.) We measured photometric redshifts from the one-dimensional spectra using a variation of the photometric redshift technique described elsewhere³⁻⁵.

The results of the analysis are photometric redshift measurements of ~ 250 objects identified in the summed direct image. Here we focus on one particular galaxy identified by the analysis at a very high redshift, which we designate as galaxy A. The J2000 coordinates of galaxy A are right ascension 12 h 36 min 27.5 s, declination $62^\circ 17' 55.4''$. Interpretation of the spectrum of galaxy A is especially straightforward because it does not overlap any other object in the dispersion direction. This means that much of the machinery of the extraction technique described above (involving de-blending overlapping spectra) does not actually come into play, although the technique is required in order to accurately determine the background sky level.

Figure 1a shows the direct image of galaxy A and its neighbours. The dispersed image of galaxy A and its neighbours, and a map of

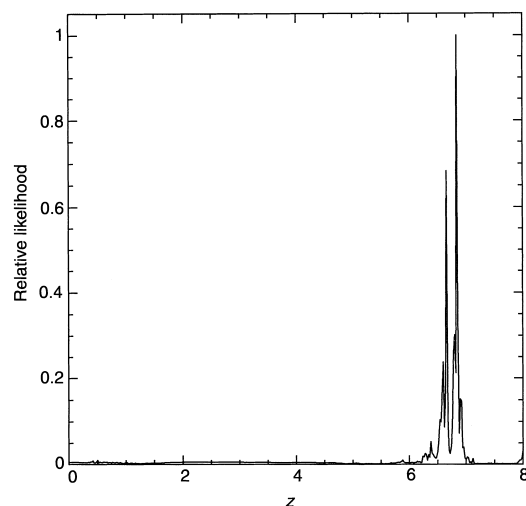


Figure 2 Redshift likelihood function of galaxy A. We measured photometric redshifts using a variation of the photometric redshift technique described previously. First, we adopted spectral templates of E/S0, Sbc, Scd, and Irr galaxies, including the effects of intrinsic and intervening neutral hydrogen absorption. (These spectral templates, which are described elsewhere^{4,5}, span rest-frame wavelengths $\lambda = 912\text{--}22,500 \text{ \AA}$.) Next we constructed the 'redshift likelihood functions' by calculating the likelihood of obtaining a measured spectrum given a modelled spectrum at an assumed redshift, maximizing with respect to galaxy spectral type and arbitrary flux normalization. Finally, we determined the maximum-likelihood photometric redshift measurements by maximizing the redshift likelihood functions with respect to redshift.

the detected emission lines, are shown in Fig. 1b and c, respectively. The dispersed image of galaxy A with best-fit model spatial profiles of all objects but galaxy A subtracted is shown in Fig. 1d. We also show the one-dimensional spectrum of galaxy A (Fig. 1e), a redshifted spectrum of a moderate-redshift ($z = 4.421$) galaxy⁶ (Fig. 1f), and the one-dimensional spectrum of galaxy A cast into 325- $\text{\AA}$ bins together with the best-fit spectrophotometric template (Fig. 1g). Figure 2 shows the redshift likelihood function of galaxy A. The spectrum of galaxy A is characterized by (1) an abrupt discontinuity at wavelength $\lambda = 9,300 \text{ \AA}$, with detectable continuum emission at wavelengths $\lambda = 9,300\text{--}9,950 \text{ \AA}$ and an absence of detectable continuum emission at wavelengths $\lambda \leq 9,300 \text{ \AA}$, and (2) an emission line at wavelength $\lambda = 9,337 \pm 6 \text{ \AA}$. The redshift likelihood function of galaxy A is characterized by a maximum at redshift $z = 6.84$ and a local maximum at redshift $z = 6.6$, with the likelihood values of the primary and secondary maxima statistically indistinguishable. The energy flux of the emission line is $(2.6 \pm 0.5) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. We interpret the abrupt discontinuity as the Lyman- α ($\text{Ly}\alpha$) decrement and the emission line as $\text{Ly}\alpha$, in which case the redshift of the galaxy determined from the emission line is $z = 6.68 \pm 0.005$.

Several lines of evidence support the validity of the spectrum extraction and the redshift identification.

(1) The integrated energy flux measured from the direct image is in excellent agreement with the integrated energy flux measured from the dispersed image. The clear magnitude of galaxy A measured from the direct image is $AB(\text{clear}) = 27.67 \pm 0.09$, and the clear magnitude of galaxy A measured from the dispersed image is $AB(\text{clear}) = 27.72 \pm 0.29$ (which is obtained by integrating the product of the spectrum and the unfiltered spectrograph and system throughput). This demonstrates that the continuum emission detected at wavelengths $\lambda \approx 9,300\text{--}9,950 \text{ \AA}$ —which is crucial for establishing the redshift of galaxy A—is of exactly the amount required to explain the direct image. The clear magnitude of the emission line alone of galaxy A measured from the dispersed image is $AB(\text{clear}) = 29.15 \pm 0.24$. This demonstrates that the emission line alone cannot explain the direct image; rather, additional continuum emission must be present, namely the continuum emission detected at wavelengths $\lambda \approx 9,300\text{--}9,950 \text{ \AA}$. It is clearly unlikely that the analysis could have missed significant continuum emission at wavelengths $\lambda < 9,300 \text{ \AA}$ —where the spectrograph and system throughput is relatively high—and instead detected spurious emission (of exactly the amount required to explain the direct image) at wavelengths $\lambda > 9,300 \text{ \AA}$.

(2) Various arguments demonstrate the reality of other faint emission lines visible in the dispersed image. For example, we applied the SExtractor program to the smoothed dispersed image to objectively identify emission lines, setting the detection threshold such that nothing was detected in the negative of the image. The resulting segmentation map is shown in Fig. 1c. Every emission line detected in the dispersed image by the SExtractor program can be attributed to a galaxy detected in the direct image. Specifically, the bottom-most emission lines arise in a very faint galaxy, designated galaxy B in Fig. 1. The spectrum of galaxy B is relatively uncomplicated, because the galaxy just barely overlaps other galaxies above and below in the dispersion direction. Only very weak continuum emission is detected from the galaxy, but at least three emission lines are clearly evident in the dispersed image (Fig. 1b and c): an emission line to the right and below the emission line of galaxy A, another emission line in the same row just over 1/2 of the way from the left-hand edge of Fig. 1, and another emission line in the same row, beyond the right-hand edge of Fig. 1. The identifications of the three emission lines are (left to right) $\text{Mg II } 2,800 \text{ \AA}$, $\text{He II } 3,203 \text{ \AA}$ and $[\text{O II}] 3,727 \text{ \AA}$ at a redshift of $z = 1.213$, which is substantiated by comparison with the ultraviolet spectrum of a starburst galaxy⁷. Furthermore, every emission line visible in Fig. 1 (including another emission line below and to the right of the emission line of galaxy A

and a 'complex' of emission lines above and to the right of the emission line of galaxy A) is robust against image processing and cosmic-ray rejection methods, and is exactly coincident with the spatial profile of a galaxy detected in the direct image, as is evident from Fig. 1d.

(3) The abrupt discontinuity is consistent with interpretation as the Ly α decrement, but is inconsistent with interpretation as other spectral breaks commonly observed in galaxy spectra. The average energy flux density measured from the dispersed image at wavelengths $\lambda = 7,250\text{--}9,250 \text{ \AA}$ is $f_{\nu}(8,250) = -0.01 \pm 0.04$, and at wavelengths $\lambda = 9,400\text{--}9,950 \text{ \AA}$ is $f_{\nu}(9,675) = 0.67 \pm 0.22 \mu\text{Jy}$, which implies a decrement $1 - f_{\nu}(8,250)/f_{\nu}(9,675) = 1.01 \pm 0.06$. The corresponding 3σ lower limit to the break amplitude is 5.9, which exceeds by a significant factor the largest measured break amplitudes of early-type galaxies (~ 2.6) and of main-sequence stars (~ 3) (ref. 8).

(4) The statistically insignificant but suggestive 'dip' in the spectrum of galaxy A at wavelengths between 9,300 and 9,500 $\text{\AA}$ matches qualitatively the expectation for stellar continuum radiation viewed through interstellar neutral hydrogen, which is expected to imprint a damped Ly α absorption feature. Specifically, the redshifted spectrum of the moderate-redshift galaxy shows a corresponding dip, as is evident from comparison of the one-dimensional spectra shown in Fig. 1e and f. A similar feature is also present in the composite spectrum of 12 high-redshift ($z \approx 3$) galaxies⁹.

(5) The redshift determined from the emission line is in excellent agreement with the redshift determined from the redshift likelihood function. Although difficult to quantify, the detection of an emission line at exactly the wavelength predicted by the photometric redshift measurement lends *a posteriori* support to the redshift identification. In summary, the spectrum of galaxy A matches exactly the spectrum expected of a galaxy of redshift approaching $z = 7$ but is unlike the spectra expected of lower-redshift galaxies.

The most striking property of galaxy A is that it is relatively bright at wavelengths longer than Ly α . Although galaxy A is exceedingly faint in any ordinary optical bandpass (due to very strong absorption by the Ly α 'forest'), its continuum energy flux density at observed-frame wavelength $\lambda \approx 9,800 \text{ \AA}$, or rest-frame wavelength $\lambda \approx 1,300 \text{ \AA}$, is $f_{\nu} \approx 1 \mu\text{Jy}$, which is comparable to the continuum energy flux densities at similar rest-frame wavelengths of galaxies identified¹⁰ at redshifts $z \approx 3$. This corresponds to an unobscured star-formation rate of $\sim 17 h^{-2} M_{\odot} \text{ yr}^{-1}$ (where $M_{\odot}$ is the solar mass) for $q_0 = 0.5$ (where q_0 is the deceleration parameter of the Universe), and $\sim 147 h^{-2} M_{\odot} \text{ yr}^{-1}$ for $q_0 = 0.0$, using the relationship¹¹ between ultraviolet luminosity and star-formation rate with a Salpeter initial mass function. Taking the redshift range searched for very-high-redshift galaxies to extend from $z = 6$ to 7, the ultraviolet luminosity density contributed by galaxy A alone is almost ten times the value measured at $z \approx 3$. If galaxy A is typical of the very-high-redshift galaxy population, then apparently it seems that very-high-redshift galaxies are luminous at rest-frame ultraviolet wavelengths, and that the unobscured cosmic star-formation rate may be substantially larger at $z \approx 7$ than at lower redshifts. $\square$

Received 24 August 1998; accepted 26 January 1999.

- Gardner, J. P. *et al.* The STIS parallel survey: introduction and first results. *Astrophys. J.* **492**, L199–L202 (1998).
- Bertin, E. & Arnouts, S. SExtractor: software for source extraction. *Astron. Astrophys. Suppl.* **117**, 393–404 (1996).
- Lanzetta, K. M., Yahil, A. & Fernández-Soto, A. Star-forming galaxies at very high redshifts. *Nature* **381**, 759–763 (1996).
- Lanzetta, K. M., Fernández-Soto, A. & Yahil, A. in *The Hubble Deep Field*, Proc. Space Telescope Science Institute 1997 May Symp. (eds Livio, M., Fall, S. M. & Madau, P.) (Cambridge Univ. Press, in the press).
- Fernández-Soto, A., Lanzetta, K. M. & Yahil, A. A new catalog of photometric redshifts in the Hubble Deep Field. *Astrophys. J.* (in the press).
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M. & Pettini, M. Lyman break galaxies at $z > 4$ and the evolution of the UV luminosity density at high redshift. *Astrophys. J.* (in the press); also as preprint astro-ph/9811399 at (<http://xxx.lanl.gov>) (23 January 1999).
- Kinney, A. L. *et al.* Template ultraviolet to near-infrared spectra of star-forming galaxies and their application to K-corrections. *Astrophys. J.* **467**, 38–60 (1998).
- Spinrad, H. *et al.* A sub-galactic pair at $z = 5.35$. *Astron. J.* **116**, 2617–2623 (1998).
- Lowenthal, J. D. *et al.* Keck spectroscopy of redshift z approximately 3 galaxies in the Hubble deep field. *Astrophys. J.* **481**, 673–688 (1997).

- Steidel, C., Giavalisco, M., Pettini, M., Dickinson, M. & Adelberger, K. Spectroscopic confirmation of a population of normal star-forming galaxies at redshifts $z > 3$. *Astrophys. J.* **462**, L17–L21 (1996).
- Madau, P., Pozzetti, L. & Dickinson, M. The star formation history of field galaxies. *Astrophys. J.* **498**, 106–116 (1998).

Acknowledgements. We thank H. Spinrad, and B. Woodgate, B. Hill and the rest of the STIS instrument team for discussions. This work was supported by NASA and the US NSF.

Correspondence should be addressed to H.-W.C. (e-mail: hchen@sbastr.ess.sunysb.edu).

Neutron scattering from magnetic excitations in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$

H. F. Fong^{*}, P. Bourges[†], Y. Sidis[‡], L. P. Regnault[‡], A. Ivanov[§], G. D. Gull^{||}, N. Koshizuka[¶] & B. Keimer^{*#}

^{*} Department of Physics, Princeton University, Princeton, New Jersey 08544, USA

[†] Laboratoire Léon Brillouin, CEA-CNRS, CE Saclay, 91191 Gif sur Yvette, France

[‡] CEA Grenoble, Département de Recherche Fondamentale sur la Matière Condensée, 38054 Grenoble cedex 9, France

[§] Institut Laue-Langevin, 156X, 38042 Grenoble Cedex 9, France

^{||} Department of Advanced Electronic Materials, School of Physics, University of New South Wales, Sydney 2052, Australia

[¶] SRL/ISTEC, 10-13, Shinonome 1-chome, Koto-ku, Tokyo 135, Japan

[#] Max-Planck-Institut für Festkörperforschung, 70569 Stuttgart, Germany

Many of the physical properties of the copper oxide high-temperature superconductors appear to defy the conventional 'one-electron' theory of metals. The development of alternative theories incorporating strong electron correlations is currently at the forefront of research in condensed-matter physics. In this context, inelastic neutron scattering can provide valuable insight into collective magnetic excitations in the copper oxide superconductors and so guide these theoretical efforts. Such measurements require large single crystals, and have hitherto been restricted to just two families of high-temperature superconductors— $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (ref. 1) and $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (ref. 2). Although the magnetic spectra of these two materials bear certain similarities, there are also important differences. In particular, a sharp resonant spin excitation dominates the spectrum in the superconducting state of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (refs 3–10), but is not observed in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (ref. 1). Here we report inelastic neutron scattering measurements of a different copper oxide system— $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ —in which we observe a magnetic resonance peak in the superconducting state. This result implies that this excitation phenomenon is a general feature of the copper oxide superconductors, so extending the empirical basis for its theoretical description.

The magnetic resonance peak observed in optimally doped $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ (superconducting transition temperature $T_c = 93 \text{ K}$) is a sharp collective mode that occurs at an energy of 40 meV and two-dimensional wavevector $Q_{\parallel} = (\pi/a, \pi/a)$, where a is the nearest-neighbour Cu–Cu distance^{3–7}. Its intensity decreases continuously with increasing temperature, and vanishes above T_c . In underdoped $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ the mode energy decreases monotonically with decreasing hole concentration^{8–10}, while the superconducting energy gap is believed to remain approximately constant. Such a collective mode has not been observed in conventional superconductors. Several microscopic mechanisms have been proposed, ranging from band structure singularities^{11–13} to antiferromagnetic spin fluctuations^{14–19} and interlayer pair tunnelling²⁰. In all of these models, the interactions that give rise to the resonant mode are the same that cause pairing of the electrons in the superconducting state, so that this phenomenon provides a direct clue to the mechanism of high-temperature superconductivity. According to another recent theory, the resonance mode is a manifestation of a new symmetry group linking antiferromagnet-

ism and superconductivity^{21,22}. In the framework of this model, the spectral weight of the resonance peak is in fact directly proportional to the superconducting condensation energy^{23,24}.

A very sensitive test of these disparate models is whether they are capable of providing detailed descriptions of both the inelastic neutron scattering (INS) results and those of angle-resolved photoemission spectroscopy (ARPES), a complementary momentum-resolved experimental technique that primarily probes single-electron excitations. By far the best ARPES data have been obtained on $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (refs 25, 26), a material for which no INS data have been available due to experimental difficulties. This situation, which has precluded a direct, quantitative comparison of the results of both techniques, is remedied by the present study.

A look at the chemical bonding in the copper oxides reveals the origin of these experimental difficulties. In $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$, for instance, the CuO_2 planes are separated by several charge reservoir layers that interact only weakly in the $\hat{c}$ -direction (perpendicular to the layers). $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ therefore cleaves easily along the CuO_2 layers, which greatly facilitates surface-sensitive techniques such as ARPES: but this same property is responsible for the lack of INS data on this compound. To take advantage of the momentum resolution of INS, the signal from a large volume of material has to be coherently superposed, which is only possible if a single crystal is available. However, because of the weak bonding along $\hat{c}$, $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ crystals typically grow as thin plates with volumes much too small for INS. The present study was performed on a comparatively large single crystal of volume $10 \times 5 \times 1.2 \text{ mm}^3$ and mosaicity (that is, the angular spread of the crystallographic axes) $\sim 1^\circ$, synthesized by the floating zone technique as described previously²⁷. Its superconducting transition temperature (measured by SQUID (superconducting quantum interference device) magnetometry) is $T_c = 91 \text{ K}$, with a width of 2 K .

The crystal was mounted in an aluminium can filled with helium exchange gas and attached to the cold finger of a closed-cycle helium refrigerator. The INS experiments were performed at the spectrometers 2T at the Laboratoire Léon Brillouin, Saclay, France, and IN8 at the Institut Laue-Langevin in Grenoble, France. Both spectrometers use neutron optics that focus the beam both parallel and perpendicular to the scattering plane, with a resulting gain in neutron flux on the sample that proved crucial for this experiment. Identical configurations (Cu(111) monochromator, PG(002) analyser, 35 meV fixed final energy) were used at both spectrometers, so that the data can be directly compared after correcting for the different neutron fluxes at both reactors. The inelastic cross-section was put on an absolute scale by a calibration against the incoherent scattering of vanadium. The comparison with $\text{YBa}_2\text{Cu}_3\text{O}_7$ (below) is also facilitated by the availability of data taken on the same spectrometer under the same experimental conditions⁶.

As our $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ crystal has a superconducting transition temperature that is similar to that of optimally doped $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$, we began our search for the resonance mode in an energy range around 40 meV. Figure 1a shows scans of the neutron intensity as a function of the in-plane wavevector $\mathbf{Q}_{\parallel}$ at a constant excitation energy of 43 meV. A featureless background, mostly due to multiphonon processes commonly observed in materials with large unit cells, is present at all temperatures. Whereas no signal can be identified above this flat background at $T = 100 \text{ K}$, a pronounced peak centred at $\mathbf{Q}_{\parallel} = (\pi/a, \pi/a)$ appears at lower temperature. The peak amplitude measured at symmetry-equivalent points in reciprocal space decreases with increasing wavevector, providing reassurance that the peak is of magnetic origin, and it compares well with the magnetic form factor of copper measured in antiferromagnetically long-range ordered copper oxides.

Indeed, the $\mathbf{Q}_{\parallel}$ -dependence of the neutron intensity at low temperatures is one of the hallmarks of the magnetic resonance peak in the superconducting state of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. Another

characteristic feature of the resonance peak is revealed by a scan of the wavevector $\mathbf{Q}_{\perp}$ perpendicular to the CuO_2 planes. As shown in Fig. 1b, the additional low-temperature intensity is sinusoidally modulated as a function of $\mathbf{Q}_{\perp}$, with a periodicity equal to the inverse of the distance of two adjacent CuO_2 planes. (Because of kinematical restrictions at low Q and phonon contamination at high Q , only one of the maxima of the sinusoid is accessible.) The modulation is a consequence of strong magnetic interactions between adjacent layers, whose microscopic nature has been the focus of much debate in the context of the magnetic resonance peak^{11–21}.

We now proceed to demonstrate the presence of the remaining signatures of the magnetic resonance peak in the $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ data. Figure 2 shows a plot of the additional low-temperature intensity as a function of the excitation energy, with the wavevector fixed at the maximum in both $\mathbf{Q}_{\parallel}$ and $\mathbf{Q}_{\perp}$. The intensity is peaked around an energy of $\sim 43 \text{ meV}$, with a width of $\sim 10\text{--}15 \text{ meV}$. Finally, Fig. 3 shows the temperature dependence of the peak

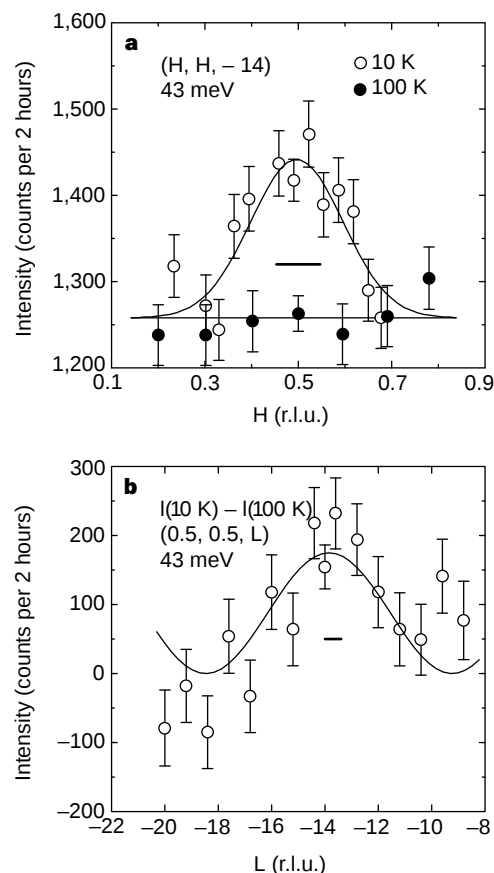


Figure 1 Constant-energy scans, at excitation energy 43 meV, of the neutron scattering intensity as a function of wavevector. The wavevectors considered are $\mathbf{Q}_{\parallel}$ (a) and $\mathbf{Q}_{\perp}$ (b), parallel and perpendicular to the CuO_2 planes, respectively. The wavevector $\mathbf{Q} = (H, K, L)$ is given in reciprocal lattice units (r.l.u.), that is, in units of the reciprocal lattice vectors $a^* \approx b^* = 1.64 \text{ \AA}^{-1}$ and $c^* = 0.20 \text{ \AA}^{-1}$. In these units, $\mathbf{Q}_{\parallel} = (0.5, 0.5)$ corresponds to the antiferromagnetic wavevector $(\pi/a, \pi/a)$. The points in **a** are raw data taken at temperatures 10 K ($< T_c$) and 100 K ($> T_c$). In **b**, the intensity at 100 K was subtracted from that at 10 K. The line in **a** is the result of a fit to a gaussian profile, while that in **b** is the function $\sin^2(\pi z_{\text{Cu}} L)$, where $z_{\text{Cu}} = 0.109$ is the distance between nearest-neighbour copper atoms in the $\hat{c}$ -direction perpendicular to the planes, expressed as a fraction of the lattice parameter $c = 30.8 \text{ \AA}$. This modulation is due to magnetic interactions between adjacent CuO_2 layers and has the same periodicity and phase as in $\text{YBa}_2\text{Cu}_3\text{O}_7$. The horizontal bars represent the experimental wavevector resolution (full-width at half-maximum).

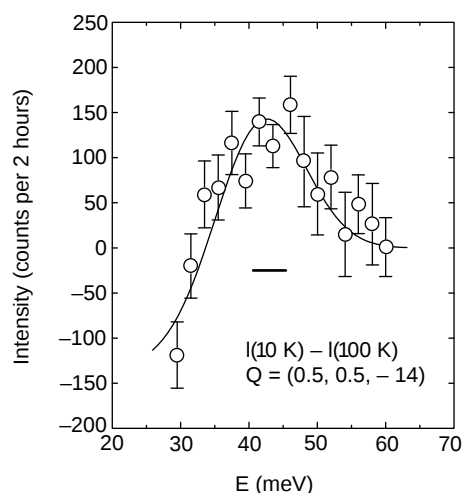


Figure 2 Difference spectrum of the neutron intensities at $T = 10\text{ K}$ ($< T_c$) and $T = 100\text{ K}$ ($> T_c$), at wavevector $\mathbf{Q} = (0.5, 0.5, -14)$. The horizontal bar represents the instrumental energy resolution; the line is a guide to the eye.

amplitude which vanishes above the superconducting transition temperature, to within the experimental uncertainty.

The magnetic excitation spectra of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ and $\text{YBa}_2\text{Cu}_3\text{O}_7$ thus exhibit an unmistakable similarity. In the superconducting state, the magnetic intensity is sharply concentrated around a single point in energy and wavevector; in the normal state, the intensity is either too broad or too weak to be observable above background. There is also no indication of magnetic intensity above background level at other energies or wavevectors. In particular, an extensive search for magnetic excitations at 10 meV —an energy below the onset of optical phonon scattering where magnetic excitations, if present, should be relatively easy to observe—has thus far been fruitless in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. In both $\text{YBa}_2\text{Cu}_3\text{O}_7$ and $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$, the resonance peak is therefore by far the most prominent feature in the magnetic excitation spectrum.

A further comparison between both compounds is made possible by a calibration of the absolute neutron cross-section against a vanadium standard. Expressed in the same units as in ref. 28, the energy-integrated spectral weight of the resonance peak at $\mathbf{Q}_\parallel = (\pi/a, \pi/a)$ is $\int d(\hbar\omega) \chi''(\mathbf{Q}, \omega) = (1.9 \pm 1) \mu_B^2$, where χ'' is the imaginary part of the dynamical spin susceptibility and $\hbar\omega$ is the excitation energy. (The large error bar is mostly due to the uncertainty in the energy width, Fig. 2.) To within the experimental error, this is identical to $1.6 \mu_B^2$ determined in $\text{YBa}_2\text{Cu}_3\text{O}_7$ (refs 7, 29). The width of the resonance in $\mathbf{Q}_\parallel$ is much larger in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ ($0.52 \text{ \AA}^{-1}$ full-width at half-maximum) than in $\text{YBa}_2\text{Cu}_3\text{O}_7$ ($0.25 \text{ \AA}^{-1}$). If averaged over the two-dimensional Brillouin zone, in addition to integrating over $\hbar\omega$, the resonance spectral weight is therefore clearly larger in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ ($\int d\mathbf{Q}_\parallel d(\hbar\omega) \chi''(\mathbf{Q}, \omega) / \int d\mathbf{Q}_\parallel = 0.23 \mu_B^2$) than in $\text{YBa}_2\text{Cu}_3\text{O}_7$ ($0.043 \mu_B^2$; ref. 29).

Such quantitative comparisons between different materials are required for a microscopic, quantitative description of the origin of the magnetic resonance peak. In the framework of the models proposed for the resonance peak, it should now be possible to relate the different $\mathbf{Q}$ -widths measured in $\text{YBa}_2\text{Cu}_3\text{O}_7$ and $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ to their respective Fermi surfaces as measured by ARPES. (The difference in the energy widths may be more difficult to interpret, as recent experiments on Zn-substituted $\text{YBa}_2\text{Cu}_3\text{O}_7$ indicate that it is very sensitive to disorder²⁹. We note that this sensitivity may account for the absence of the resonance peak in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, where disorder due to Sr donors is significant. The $\mathbf{Q}$ -width, on the other hand, is not strongly affected by impurities.) There have also been attempts to relate more subtle features of the ARPES spectra, such as the well-known ‘dip’ in the spectral

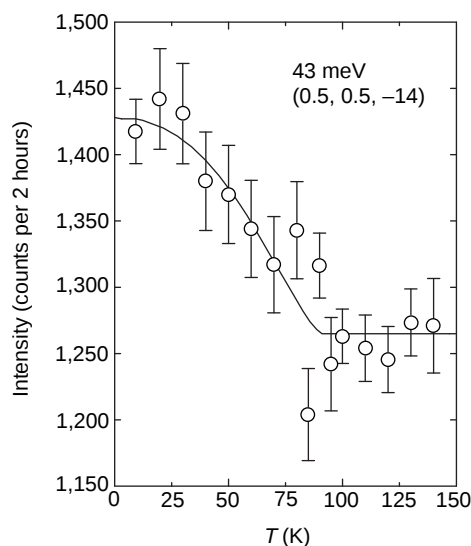


Figure 3 Temperature dependence of the neutron intensity at energy 43 meV and wavevector $\mathbf{Q} = (0.5, 0.5, -14)$. The intensity falls to background level above $T_c = 91\text{ K}$ (Fig. 1). The line is a guide to the eye.

function, to the collective spin dynamics^{30,31}. These features are only clearly apparent in the ARPES work on $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$, and our INS data will now allow such arguments to be refined and put on a quantitative footing. The same applies to recent models that propose a direct relation between the spectral weight of the resonance peak and the superconducting condensation energy^{23,24}.

Our study opens the way for a variety of further neutron experiments, in particular in the overdoped regime which is easily accessible in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ over a wide range of hole concentrations. It also leaves open questions that can only be answered by neutron-scattering work on other families of high-temperature superconductors. For instance, as both $\text{YBa}_2\text{Cu}_3\text{O}_7$ and $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ are bilayer materials, the present study does not provide further insight into the role of interlayer interactions in the resonance peak. Most importantly, observation of the resonance peak in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ rules out the possibility that this phenomenon is due to a conspiracy of structural or chemical parameters peculiar to $\text{YBa}_2\text{Cu}_3\text{O}_7$. It is in fact an intrinsic feature of the copper oxides, and an explanation of this feature needs to be an integral part of any theory of high-temperature superconductivity. □

Received 9 December 1998; accepted 12 February 1999.

- Kastner, M. A., Birgeneau, R. J., Shirane, G. & Endoh, Y. Magnetic, transport, and optical properties of monolayer copper oxides. *Rev. Mod. Phys.* **70**, 897–928 (1998).
- Tranquada, J. M., Shirane, G., Keimer, B., Shamoto, S. & Sato, M. Neutron scattering study of magnetic excitations in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. *Phys. Rev. B* **40**, 4503–4516 (1989).
- Rossat-Mignod, J. *et al.* Neutron scattering study of the $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ system. *Physics C* **185–189**, 86–92 (1991).
- Mook, H. A., Yethiraj, M., Aeppli, G., Mason, T. E. & Armstrong, T. Polarized neutron determination of the magnetic excitations in $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. Lett.* **70**, 3490–3493 (1993).
- Fong, H. F. *et al.* Phonon and magnetic neutron scattering at 41 meV in $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. Lett.* **75**, 316–319 (1995).
- Bourges, P., Regnault, L. P., Sidis, Y. & Vettier, C. Inelastic-neutron-scattering study of antiferromagnetic fluctuations in $\text{YBa}_2\text{Cu}_3\text{O}_{6.97}$. *Phys. Rev. B* **53**, 876–885 (1996).
- Fong, H. F., Keimer, B., Reznik, D., Dogan, F. & Aksay, I. A. Polarized and unpolarized neutron-scattering study of the dynamical spin susceptibility of $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. B* **54**, 6708–6720 (1996).
- Fong, H. F., Keimer, B., Millius, D. L. & Aksay, I. A. Superconductivity-induced anomalies in the spin excitation spectra of underdoped $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. *Phys. Rev. Lett.* **78**, 713–716 (1997).
- Dai, P., Yethiraj, M., Mook, H. A., Lindemer, T. B. & Dogan, F. Magnetic dynamics in underdoped $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$: Direct observations of a superconducting gap. *Phys. Rev. Lett.* **77**, 5425–5428 (1996).
- Bourges, P. *et al.* Shifting of the magnetic-resonance peak to lower energy in the superconducting state of underdoped $\text{YBa}_2\text{Cu}_3\text{O}_{6.8}$. *Europhys. Lett.* **38**, 313–318 (1997).
- Bulut, N. & Scalapino, D. J. Neutron scattering from a collective spin fluctuation mode in a CuO_2 bilayer. *Phys. Rev. B* **53**, 5149–5152 (1996).
- Blumberg, G., Stojkovic, B. P. & Klein, M. V. Antiferromagnetic excitations and van Hove singularities in $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. *Phys. Rev. B* **52**, 15741–15744 (1995).
- Abrikosov, A. A. Neutron peak in the extended-saddle-point model of high-temperature superconductors. *Phys. Rev. B* **57**, 8656–8661 (1998).
- Liu, D. Z., Zha, Y. & Levin, K. Theory of neutron scattering in the normal and superconducting states of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$. *Phys. Rev. Lett.* **75**, 4130–4133 (1995).

15. Mazin, I. I. & Yakovenko, V. M. Neutron scattering and superconducting order parameter in $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. Lett.* **75**, 4134–4137 (1995).
16. Onufrieva, F. & Rossat-Mignod, J. Theory of spin dynamics in high- T_c copper oxide superconductors: Application to neutron scattering. *Phys. Rev. B* **52**, 7572–7603 (1995).
17. Onufrieva, F. Evidence for $d_{x^2-y^2}$ symmetry of the superconducting order parameter in YBCO from neutron-scattering data. *Physica C* **251**, 348–354 (1995).
18. Millis, A. J. & Monien, H. Bilayer coupling in the yttrium–barium family of high-temperature superconductors. *Phys. Rev. B* **54**, 16172–16178 (1996).
19. Morr, D. K. & Pines, D. The resonance peak in cuprate superconductors. *Phys. Rev. Lett.* **81**, 1086–1089 (1998).
20. Yin, L., Chakravarty, S. & Anderson, P. W. The neutron peak in the interlayer tunneling model of high temperature superconductors. *Phys. Rev. Lett.* **78**, 3559–3562 (1997).
21. Zhang, S. C. A unified theory based on $\text{SO}(5)$ symmetry of superconductivity and antiferromagnetism. *Science* **275**, 1089–1096 (1997).
22. Demler, E., Kohno, H. & Zhang, S. C. π -excitation of the t - J model. *Phys. Rev. B* **58**, 5719–5730 (1998).
23. Scalapino, D. J. & White, S. R. The superconducting condensation energy and an antiferromagnetic exchange based pairing mechanism. *Phys. Rev. B* **58**, 8222–8224 (1998).
24. Demler, E. & Zhang, S. C. Quantitative determination of a microscopic mechanism of high T_c superconductivity. *Nature* **396**, 733–735 (1998).
25. Norman, M. R. *et al.* Destruction of the Fermi surface in underdoped high- T_c superconductors. *Nature* **392**, 157–160 (1998).
26. Shen, Z. X. *et al.* Temperature-induced momentum-dependent spectral weight transfer in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Science* **280**, 259–262 (1998).
27. Gu, G. D., Takamuku, K., Koshizuka, N. & Tanaka, S. Large single crystal Bi-2212 along c -axis prepared by floating zone method. *J. Cryst. Growth* **130**, 325–329 (1998).
28. Bourges, P. *et al.* High-energy spin excitations in $\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$. *Phys. Rev. B* **56**, R11439–R11442 (1997).
29. Fong, H. F. *et al.* Effect on nonmagnetic impurities on the magnetic resonance peak in $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. Lett.* **82**, 1939–1942 (1999).
30. Norman, M. R. & Ding, H. Collective modes and the superconducting-state spectral function of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$. *Phys. Rev. B* **57**, R11089–R11092 (1998).
31. Shen, Z. X. & Schrieffer, J. R. Momentum, temperature, and doping dependence of photoemission lineshape and implications for the nature of the pairing potential in high- T_c superconducting materials. *Phys. Rev. Lett.* **78**, 1771–1774 (1997).

Acknowledgements. We acknowledge discussions with A. H. Moudden. The work at Princeton University was supported by the MRSEC programme of the National Science Foundation, and by the Packard and Sloan Foundations.

Correspondence and requests for materials should be addressed to B.K. (e-mail: keimer@pupppg.princeton.edu).

Surface-aligned reaction of photogenerated oxygen atoms with carbon monoxide targets

C. Emil Tripa & John T. Yates Jr

Department of Chemistry, Surface Science Center, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Atomic and molecular species generated by the photolysis of aligned molecules adsorbed on crystalline solids tend to move preferentially in particular directions relative to the crystal surface¹. This behaviour results in surface-aligned photo-reaction^{1,2} if the photogenerated species is directed towards, and reacts with, adsorbed and aligned target molecules. Previously, geometrical directionality has been inferred from the reaction product yield, the angular distribution and/or the kinetic and internal energy distributions of departing photochemically

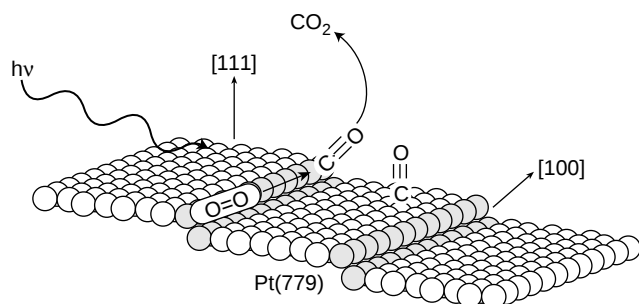


Figure 1 Schematic picture of the preferred CO photo-oxidation on a stepped Pt single-crystal surface.

produced species^{3–6}. Here we report measurements of the relative rate of the reaction between oxygen atoms (photogenerated from adsorbed and surface-aligned molecular oxygen) and carbon monoxide molecules adsorbed on either the step or the terrace sites of platinum single crystals. By using isotopically distinct carbon monoxide molecules, we are able to show that the oxidation rate at step sites is twice the oxidation rate at terrace sites. This observation suggests that the motion of the photogenerated oxygen atoms is aligned along the step edge, so that the atoms are ‘aimed’ at carbon monoxide molecules adsorbed at step sites.

Figure 1 gives a schematic illustration of the surface-aligned photochemical (SAP) reaction we have investigated. Our experiments were carried out on two stepped Pt single crystals. The Pt(335) surface ($\text{Pt(s)}-[100] \times 4[111]$) contains (100) step sites of single-atom height, separated by (111) terrace sites that are four atoms in width. The Pt(779) surface ($\text{Pt(s)}-[100] \times 8[111]$) also contains single-atom-height (100) steps but has terraces of double the width of Pt(335). Oxygen molecules are preferentially trapped at step sites from a mobile precursor state on adsorption at 85–90 K (ref. 7) and orientated with their O–O axes parallel to the step edge^{8,9}. Thermal dissociation of chemisorbed O_2 does not occur on Pt(335) at temperatures below 140 K (refs 10, 11).

Atomically clean Pt surfaces were exposed at 88 K to O_2 doses from an absolutely calibrated and collimated beam doser to achieve reproducibly $\sim 13\%$ of a monolayer O_2 coverage on Pt(335) and $\sim 7\%$ O_2 coverage on Pt(779) (with 50% filling of the step sites in each case). $^{13}\text{C}^{18}\text{O}$ (99% ^{13}C and 95% ^{18}O isotopic purity) was then admitted to the surface at 88 K to fill the step sites, and the absence of CO molecules on the terrace sites was confirmed by observation

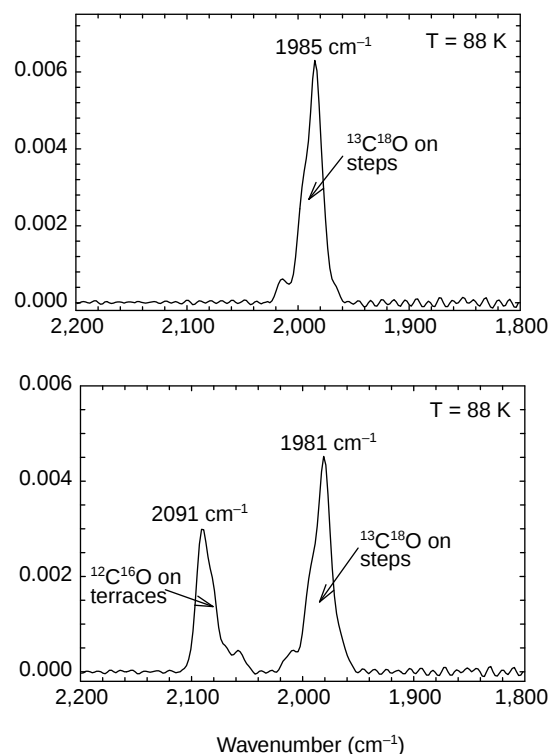


Figure 2 Selective filling of the adsorption sites with CO isotopomers on the stepped Pt(335) surface. This process is monitored by reflection-absorption infrared spectroscopy (RAIR). The step sites are first saturated with O_2 (not visible in this IR spectrum) and then with $^{13}\text{C}^{18}\text{O}$ ($1,985\text{ cm}^{-1}$), as shown in the upper spectrum. After the step sites are blocked in this fashion, the $^{12}\text{C}^{16}\text{O}$ isotopomer ($2,091\text{ cm}^{-1}$) can only occupy the terrace sites (the lower spectrum). A $\sim 20\%$ intensity sharing effect is observed, judging from the decrease in step- $^{13}\text{C}^{18}\text{O}$ absorbance as terrace- $^{12}\text{C}^{16}\text{O}$ is added.

of the C–O stretching-mode frequency by reflection infrared spectroscopy^{12,13}. Isotopomer $^{12}\text{C}^{16}\text{O}$ (98.9% isotopic purity) was then placed on the terrace sites so that equal numbers of the two CO isotopomers were present on the two types of sites. This equality of isotopomer coverage was measured by separate measurements of isotopic CO_2 production using temperature-programmed reaction. A typical site-filling sequence (Fig. 2) shows that step sites containing chemisorbed O_2 then become selectively saturated with $^{13}\text{C}^{18}\text{O}$, and that no CO site exchange occurs upon filling the terrace sites at 88 K with $^{12}\text{C}^{16}\text{O}$ since terrace-bound $^{13}\text{C}^{18}\text{O}$ would exhibit a frequency near $2,000\text{ cm}^{-1}$. The intensity sharing effect^{12,14,15} due to vibrational coupling between the chemisorbed CO on the two types of sites was minimized by the use of the heaviest and lightest isotopomers of CO for these experiments.

The filling of step and terrace sites with isotopically distinct CO molecules allows us to monitor separately the photo-oxidation of CO from each site (Fig. 3). For both crystals, the rate of loss of CO from the step sites exceeds the rate of loss from the terrace sites. First-order fits to the data, involving six sequential infrared measurements, show that the rate of CO depletion from the step

sites is 2.1 ± 0.2 times greater than the rate from the terrace sites. This factor of about 2 in relative rate of depletion of step-CO compared to terrace-CO is obtained from measurements on both Pt crystals and thus is independent of terrace width. The reaction ends when the chemisorbed O_2 on the step sites has been depleted. Photodissociation of O_2 results mainly in the desorption of O_2 from the Pt crystal or the capture of atomic O by the surface (C.E.T. and J.T.Y., manuscript in preparation). As indicated by the relatively small decrease ($<5\%$) in CO coverage and in agreement with our earlier findings¹⁶, CO_2 production is only of minor importance. In experiments on both Pt crystals, we also observe the appearance and intensification of a small absorption peak near $2,060\text{ cm}^{-1}$ with increasing irradiation time (Fig. 3), which is attributed to the C–O stretching mode of $^{12}\text{C}^{16}\text{O}$ adsorbed on step sites. The appearance of this peak is caused by the transfer of $^{12}\text{C}^{16}\text{O}$ from terrace sites to one of the increasing number of empty step sites created by the photochemical processes taking place on the Pt crystals.

Instead of following the disappearance of the reactant CO, the relative rate of reaction as a function of CO on a particular type of adsorption site can also be inferred from differences in the rate of product formation. Products can be differentiated because the reaction between photochemically produced ^{16}O atoms and $^{13}\text{C}^{18}\text{O}$ at the step sites and $^{12}\text{C}^{18}\text{O}$ at the terrace sites should yield $^{13}\text{C}^{16}\text{O}^{18}\text{O}$ and $^{12}\text{C}^{16}\text{O}^{18}\text{O}$, respectively. (In this experiment, $^{12}\text{C}^{18}\text{O}$, rather than $^{12}\text{C}^{16}\text{O}$, was placed on terrace sites to allow an unambiguous identification of CO_2 as being the result of the interaction between adsorbed CO and photogenerated O atoms. We analysed the random flux of CO_2 produced through the photo-oxidation of

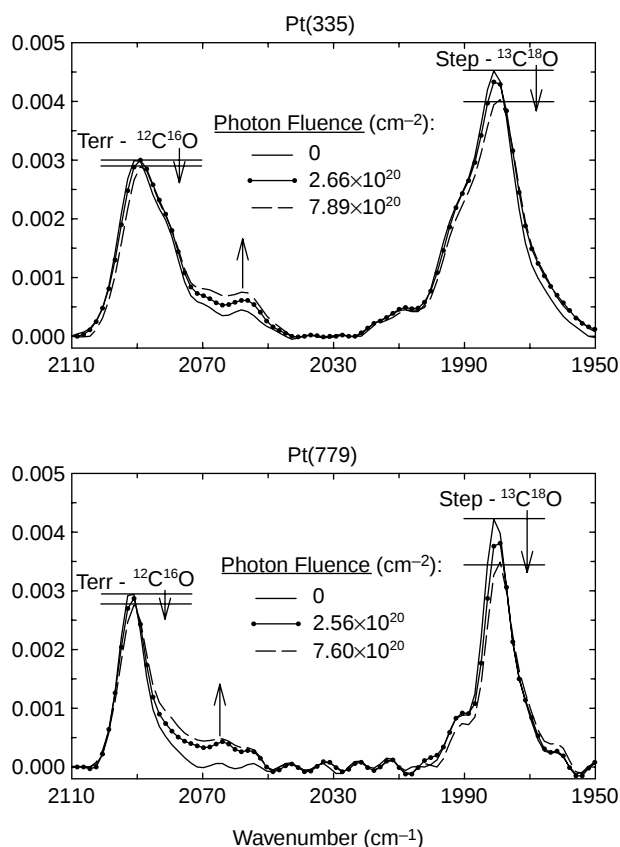


Figure 3 Photo-oxidation of CO by O_2 on a Pt(335) single crystal (upper panel) and a Pt(779) single crystal (lower panel) as observed by RAIR spectroscopy at different irradiation times. (The irradiation times correspond to the photon fluences shown by the key in each panel.) The more pronounced decrease of the $^{13}\text{C}^{18}\text{O}$ peak, relative to that of the $^{12}\text{C}^{16}\text{O}$ peak, reflects the preferential consumption of CO on step sites over CO on terrace sites. The peak developing at $2,060\text{ cm}^{-1}$ corresponds to $^{12}\text{C}^{16}\text{O}$ undergoing site-exchange to occupy step sites vacated as a result of the photo-induced reaction between O_2 and $^{13}\text{C}^{18}\text{O}$. Photo-oxidation was induced by radiation in the range 3.87–4.77 eV from a high-pressure Hg arc lamp. IR radiation was filtered out and the photon flux ($4.5 \times 10^{17}\text{ cm}^{-2}\text{ s}^{-1}$) was measured with a photodiode having a quantum efficiency of $\sim 85\%$ (ref. 16). In agreement with observations in similar studies^{16,18}, the small radiation heating effect of 4 K initiated neither reactions between the adsorbed molecules nor the desorption of O_2 or CO.

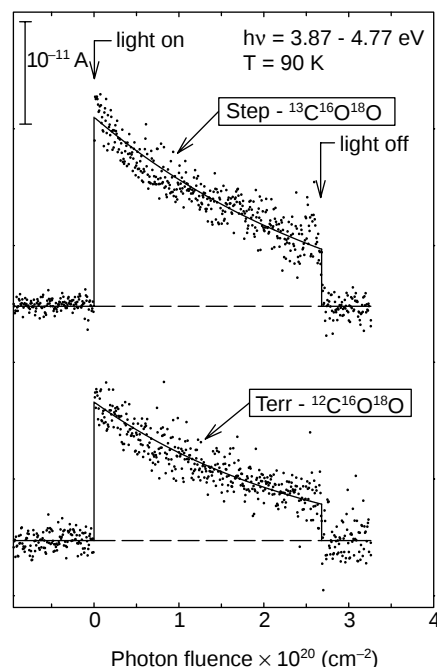


Figure 4 Photogeneration of gas-phase CO_2 from $^{16}\text{O}_2$ and CO adsorbed on Pt(335). Shown is the mass-spectrometer CO_2 signal, which is directly proportional to the rate of CO_2 production, as a function of photon fluence. (The mass spectrometer signals are measured as currents and given in units of amperes.) $^{13}\text{C}^{18}\text{O}$ and $^{12}\text{C}^{18}\text{O}$ were adsorbed on the step and terrace sites, respectively. The crystal surface was exposed to UV radiation for a period of 600 s, with switching on and off of the radiation indicated by arrows. The use of labelled CO isotopes not only allowed us to distinguish between CO on step and terrace sites, but also minimized interference from background gases. The CO_2 signals are fitted by first-order exponentials (shown as solid lines), reflecting the depletion of O_2 over the course of an experiment. The observed kinetics are expected to be of higher order because CO is depleted as well. However, the extent of CO depletion is so small that we are unable to detect deviations from first-order kinetics.

CO on a Pt(335) single crystal with a quadrupole mass spectrometer. As illustrated by Fig. 4, the initial rate of CO₂ formation from CO on step sites deduced from these data is 1.5 ± 0.2 times greater than that of CO₂ formation from CO on terrace sites. This value for the ratio of the photochemical oxidation rates is somewhat smaller than that inferred from the CO consumption measurements, but does confirm the preferential production of CO₂ at step sites.

Figure 5 schematically indicates, in a two-dimensional model, the geometrical factors involved in the SAP effect. As shown in the figure, O₂ molecules are chemisorbed by bonding to two Pt atoms on the Pt(100) step^{8,9}. A CO molecule is located in this model on an adjacent Pt atom on the Pt(100) step. We assume that the reaction cross-section between a photochemically ejected O atom and CO is $\sigma_{rx} = 17.6 \text{ \AA}^2$, as found for isotopic exchange of photochemically produced O atoms (from photolysis of NO₂) with a CO(g) molecule¹⁷. This value gives an impact parameter $b_{rx} = (\sigma_{rx}/\pi)^{1/2}$ of $2.37 \text{ \AA}$, which may be used to estimate the angular range ϕ_{step} of O-atom ejection leading to the oxidation of CO at step sites. In the case of O₂ and CO occupying adjacent Pt step atoms, photoinduced O atom ejection from the adsorbed O₂ molecule at angles not exceeding $\phi_{step} = 86^\circ$ will result in the oxidation of the adjacent CO molecule. Ejection at larger angles will bring the photogenerated O atom into the 'terrace reaction zone' and result in the oxidation of CO molecules occupying terrace sites. As shown in Fig. 5, the angular range for this reaction is $\phi_{terrace} = 2 \times 47^\circ = 94^\circ$. The ratio of the two angular ranges is close to unity ($\phi_{terrace}/\phi_{step} = 1.1$), implying that in the absence of surface alignment effects the oxidation of CO at terrace sites would be marginally favoured over the oxidation of CO at step sites. (This assumes that the cross-sections of the oxygen atom reaction with a CO molecule located on either the step or the terrace are essentially identical. The small difference of only 13 cm^{-1} , in the zero coverage limit, between the stretching-mode frequencies of ¹²C¹⁶O adsorbed on step and terrace sites, respectively¹⁴, supports this assumption.) We note that an increase in the average distance between O₂ and CO along the Pt(100) step would favour the oxidation of CO on terrace sites even more strongly in the model of Fig. 5. Our measurements indicate, in contrast, an enhancement in the rate of photo-oxidation at step sites by a factor of about two, thus supporting the role of surface alignment in controlling the relative motion of the photogenerated O atom towards its CO target molecule. The observation that doubling of the terrace width in going from Pt(335) to Pt(779)

does not affect the magnitude of the enhancement suggests that the relative rate of the photo-oxidation reaction is not affected by CO availability: most of the reactive O atoms generated will find a CO molecule on either step or terrace sites, with the observed enhancement reflecting the efficiency of the crystal surface in optimizing the alignment of the two reactants along step sites. □

Received 25 September 1998; accepted 8 February 1999.

- Polanyi, J. C. & Zeiri, Y. in *Laser Spectroscopy and Photochemistry on Metal Surfaces* Ch. 26 (eds Dai, H. L. & Ho, W.) p. 1242 (Adv. Ser. Phys. Chem. Vol. 5, part II, World Scientific, Singapore, 1995).
- Bourdon, E. B. D. *et al.* Photodissociation, photoreaction and photodesorption of adsorbed species. *Faraday Discuss. Chem. Soc.* **82**, 343–358 (1986).
- Polanyi, J. C. & Rieley, H. in *Dynamics of Gas-Surface Interactions* Ch. 8 (eds Rettner, C. T. & Ashfold, M. N. R.) p. 329 (Adv. in Gas-Phase Photochemistry and Kinetics ser., Royal Society of Chemistry, Cambridge, 1991).
- Yamanaka, T., Inoue, Y. & Matsushima, T. An aligned oxidation of carbon monoxide induced by 193 nm irradiation on platinum(113). *J. Chem. Phys. Lett.* **264**, 180–185 (1997).
- Garrett, S. J., Heyd, D. V. & Polanyi, J. C. Photochemistry of adsorbed molecules. XVIII. Photodissociation and exchange reaction in CH₃Br/MgO(001) at 193 nm. *J. Chem. Phys.* **106**, 7847–7855 (1997).
- Dixon-Warren, St. J., Jensen, E. T. & Polanyi, J. C. Photochemistry of adsorbed molecules. XI. Charge-transfer photodissociation and photoreaction in chloromethanes on Ag(111). *J. Chem. Phys.* **98**, 5938–5953 (1993).
- Winkler, A., Guo, X., Siddiqui, H. R., Hagans, P. L. & Yates, J. T. Jr Kinetics and energetics of oxygen adsorption on Pt(111) and Pt(112)—a comparison of flat and stepped surfaces. *Surf. Sci.* **201**, 419–443 (1988).
- Yamanaka, T., Matsushima, T., Tanaka, S. & Kamada, M. Orientation of oxygen admolecules on a stepped platinum (113) surface. *Surf. Sci.* **349**, 119–128 (1996).
- Stipe, B. C. *et al.* Single-molecule dissociation by tunneling electrons. *Phys. Rev. Lett.* **78**, 4410–4413 (1997).
- Rar, A. & Matsushima, T. Desorption and dissociation of oxygen admolecules on a stepped platinum (533) surface. *Surf. Sci.* **318**, 89–96 (1994).
- Wang, H., Tobin, R. G., Lambert, D. K., DiMaggio, C. L. & Fisher, G. B. Adsorption and dissociation of oxygen on Pt(335). *Surf. Sci.* **372**, 267–278 (1997).
- Xu, J., Henriksen, P. & Yates, J. T. Jr Direct spectroscopic observation of the reactive catalytic site for CO oxidation on Pt(335). *J. Chem. Phys.* **97**, 5250–5252 (1992).
- Jansch, H. J., Xu, J. & Yates, J. T. Jr Electron stimulated surface migration of CO on Pt(335). First spectroscopic evidence for a new phenomenon. *J. Chem. Phys.* **99**, 721–724 (1993).
- Hayden, B. E., Kretschmar, K., Bradshaw, A. M. & Greenler, R. G. An infrared study of the adsorption of CO on a stepped platinum surface. *Surf. Sci.* **149**, 394–406 (1985).
- Leibslle, F. M., Sorbello, R. S. & Greenler, R. G. Coupled harmonic oscillator models of carbon monoxide adsorbed on stepped, platinum surfaces. *Surf. Sci.* **179**, 101–118 (1987).
- Trippa, C. E., Arumainayagam, C. R. & Yates, J. T. Jr Kinetics measurements of CO photo-oxidation on Pt(111). *J. Chem. Phys.* **105**, 1691–1696 (1996).
- Jaffe, S. & Klein, E. S. Isotopic exchange reactions of atomic oxygen produced by the photolysis of NO₂ at 3660 Å. *Trans. Faraday Soc.* **62**, 3135–3141 (1966).
- Mieher, W. D. & Ho, W. Bimolecular surface photochemistry: Mechanisms of CO oxidation on Pt(111) at 85 K. *J. Chem. Phys.* **99**, 9279–9295 (1993).

Acknowledgements. We thank D. Heyd and J. Polanyi, for reading this manuscript, and P. E. Siska for helping us with the gas-phase kinetics literature. This work was supported by the US Department of Energy/Office of Basic Energy Sciences.

Correspondence and requests for materials should be addressed to J.T.Y. (e-mail: jyates@vms.cis.pitt.edu).

Template-imprinted nanostructured surfaces for protein recognition

Huaiqiu Shi*, Wei-Bor Tsai*, Michael D. Garrison*, Sandro Ferrari† & Buddy D. Ratner*‡

* Department of Bioengineering, University of Washington, Box 351720, Seattle, Washington 98195, USA

† Ingengneria Dei Materiali, Università Degli Studi Di Trento, Italy

‡ Department of Chemical Engineering, University of Washington, Box 351750, Seattle, Washington 98195, USA

Synthetic materials capable of selectively recognizing proteins are important in separations¹, biosensors² and the development of biomedical materials^{3–5}. The technique of molecular imprinting creates specific recognition sites in polymers by using template molecules^{6–9}. Molecular recognition is attributed to binding sites that complement molecules in size, shape and chemical functionality¹⁰. But attempts to imprint proteins have met with only limited success^{11–15}. Here we report a method for imprinting surfaces with protein-recognition sites. We use radio-frequency

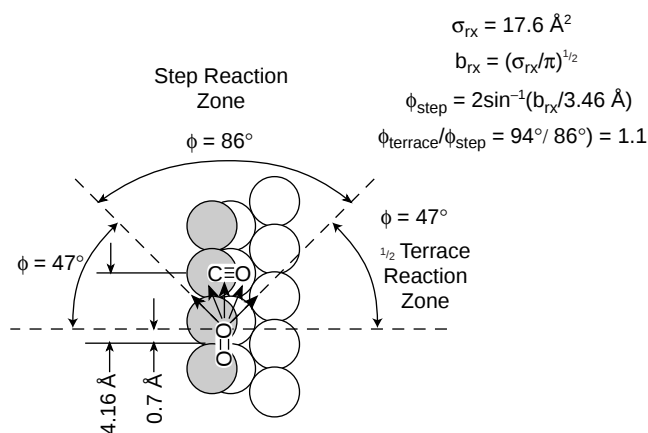


Figure 5 A model for SAP based on gas-phase kinetics. The impact parameter (b_{rx}) derived from the gas-phase kinetics of the $^{18}\text{O} + \text{C}^{16}\text{O} \rightarrow \text{C}^{18}\text{O} + ^{16}\text{O}$ isotopic exchange process (ref. 17) is used to estimate the angular dimensions of the two reaction zones. The O₂ and CO molecules are located on the nearest-neighbour step sites in this model and the separation between them is estimated from the Pt lattice constant and the O–O bond distance in the O₂ molecule adsorbed on step Pt sites (ref. 8).

glow-discharge plasma deposition to form polymeric thin films¹⁶ around proteins coated with disaccharide molecules. The disaccharides become covalently attached to the polymer film, creating polysaccharide-like cavities that exhibit highly selective recognition for a variety of template proteins, including albumin, immunoglobulin G, lysozyme, ribonuclease and streptavidin. Direct imaging of template recognition is achieved by patterning a surface at the micrometre scale with imprinted regions.

Our approach to template imprinting of proteins is outlined in Fig. 1a. Mica is used as the adsorption substrate for template proteins because it is atomically flat. Thus, the imprint surface reflects only the topographic features of the template molecules. Moreover, the hydrophilic, negatively charged surface of mica minimizes denaturation of adsorbed proteins¹⁷. Disaccharide is coated onto proteins, because multiple hydrogen bonds will be formed between the hydroxyl groups on sugars and the surface polar residues of proteins during dehydration¹⁸. The 'sugar shell' around

the protein molecule thus inhibits drying-induced denaturation or plasma-induced degradation, so that surface structural information from the template protein that is essential for recognition can be transferred to the imprinted cavity.

In radio-frequency glow-discharge (RFGD) plasma deposition of fluoropolymers such as hexafluoropropylene (C_3F_6), reactive species link with each other as well as to organic moieties on the surface. This yields a smooth, conformal film with good mechanical and chemical stability. By optimizing the conditions of plasma reaction, we covalently incorporated the coated sugars into the plasma film while minimizing the damage to the underlying protein. Mounting the plasma film on a solid support facilitates the detachment of the mica. Elution with a basic solution of the exposed mica-plasma film interface results in the decomposition and removal of the embedded protein. Protein-imprinted nanocavities are thus created that have a shape complementary to the protein and, we suggest, also have precisely positioned hydroxyl groups from the immobilized sugars.

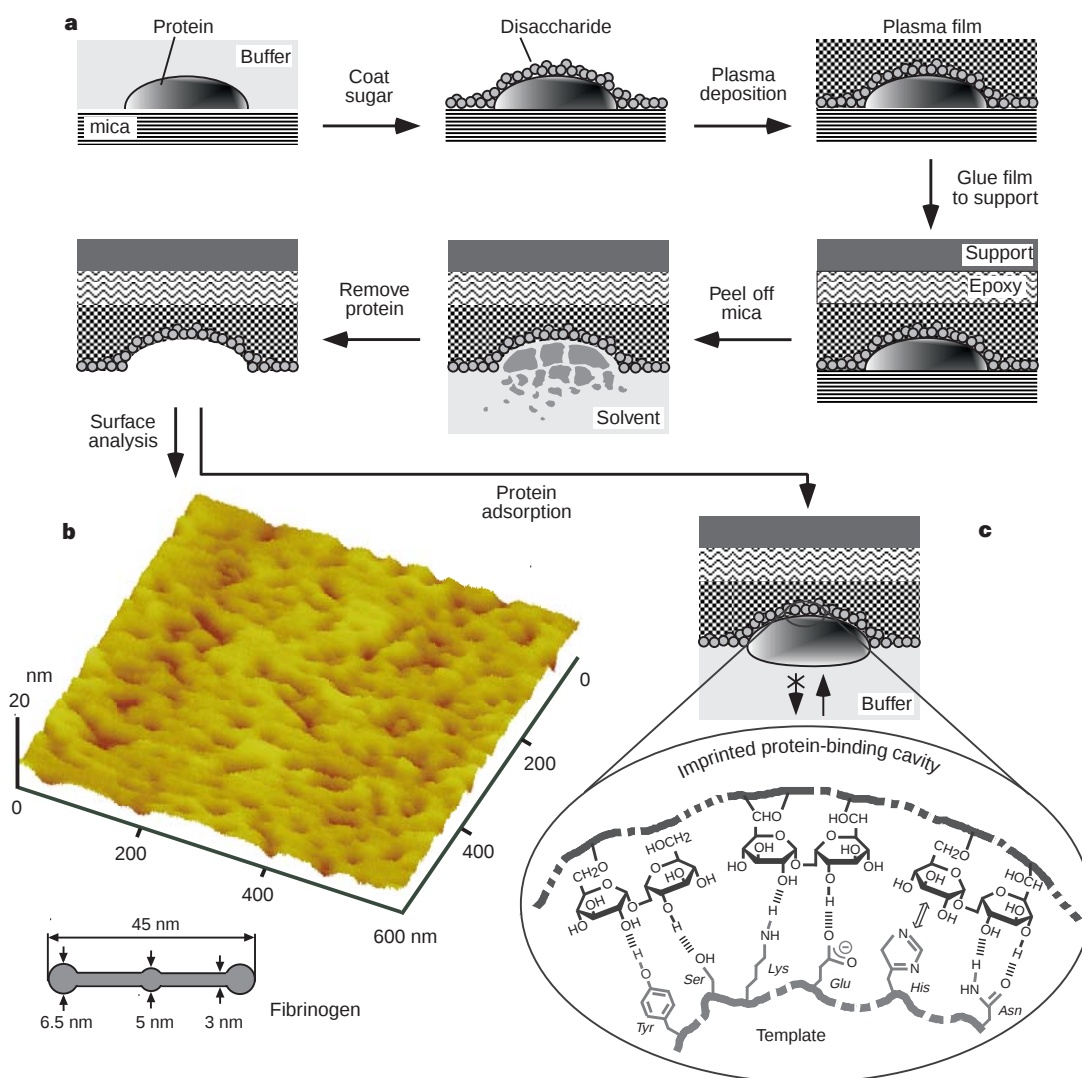


Figure 1 Protocol for template imprinting of proteins. (See refs 16 and 29 for details.) **a**, Template protein was adsorbed onto a freshly cleaved mica in citrated phosphate-buffered saline (CPBS), pH 7.4. A 1–10 mM solution of disaccharide was spin-cast to form a 10–50 Å sugar overlayer. The sample was put into the in-glow region of a 13.56 MHz RFGD reactor. Plasma deposition of C_3F_6 was conducted at 150 mtorr and 20 W for 3–6 min, forming a 10–30 nm fluoropolymer thin film. The resulting plasma film was fixed to a glass coverslip using epoxy resin and oven-cured. Mica was peeled off and the sample was soaked in a NaOH/NaClO (0.5/1.0%) solution for 0.5–2 h for dissolution and extraction of

protein. A nanopit with a shape complementary to the protein was created on the imprint surface. **b**, A tapping mode AFM image of the surface of a fibrinogen imprint, together with a drawing of fibrinogen. **c**, Mechanisms for the specific protein recognition of template-imprinted surfaces. A nanocavity-bound template protein is prevented from exchange with other protein molecules in the solution because of steric hindrance and an overall strong interaction; the latter is due to many cooperative weak interactions, involving hydrogen bonds, van der Waals forces and hydrophobic interactions for example.

We propose that these nanocavities can specifically recognize their template proteins. Hydrogen bonding is essential for molecular recognition of biological systems in aqueous media^{19–21}. Hydrogen-bond-based molecular imprints also show specific recognition for small molecules and proteins in aqueous solutions^{15,22}. When a template protein molecule enters a complementary pit, a multitude of simultaneous hydrogen bonds can be formed between the oriented polar groups within the pit and the protein surface polar residues (Fig. 1c). These cooperative, multivalent interactions lead

to a significantly increased overall protein-binding affinity. Because the directional nature of a hydrogen bond demands a high degree of structural complementarity, the number of hydrogen bonds between a pit and its non-template protein is small. Consequently, their binding strength is low and nonspecific protein binding is inhibited. Furthermore, van der Waals forces may also contribute to specific recognition because of steric complementarity between the template protein and the imprinted pit.

We imprinted three abundant blood proteins, bovine serum

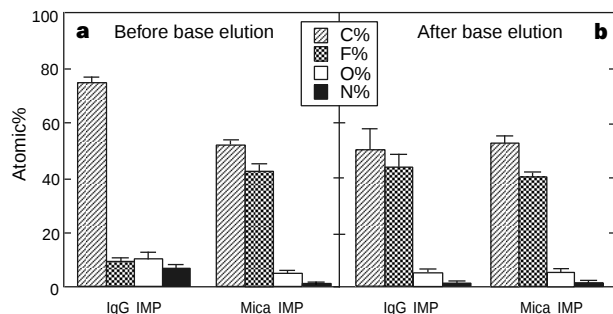
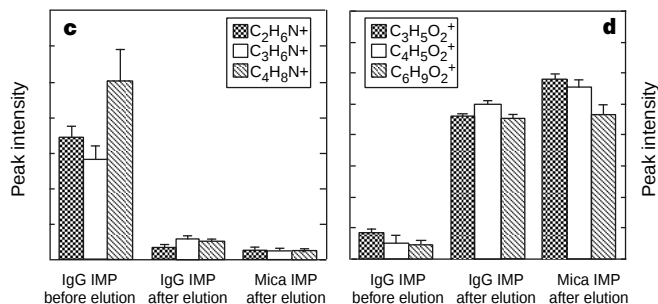


Figure 2 Surface elemental compositions measured by ESCA. **a**, IgG imprint and mica imprint control before elution with NaOH/NaClO. **b**, After elution. IMP, imprint. The insignificant change in the mica imprint after base elution is indicative of the stability of imprint surfaces when exposed to NaOH/NaClO. The IgG imprint exhibits a large decrease in N, C and O concentration after the elution by base and becomes similar to the mica imprint, suggesting that template proteins



have been removed. **c**, **d**, TOF-SIMS measurements of the surface chemistry, showing peaks characteristic of proteins (**c**) and sugars (**d**). All peaks were normalized to total counts. The IgG imprint is dominated by protein-associated peaks before base elution, but becomes rich in sugar-associated peaks afterwards, confirming that template proteins have been extracted by base elution.

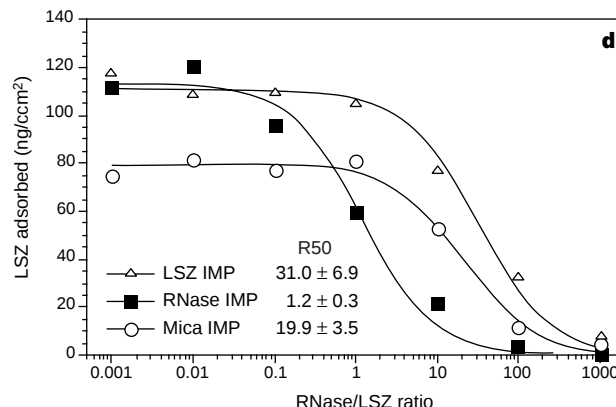
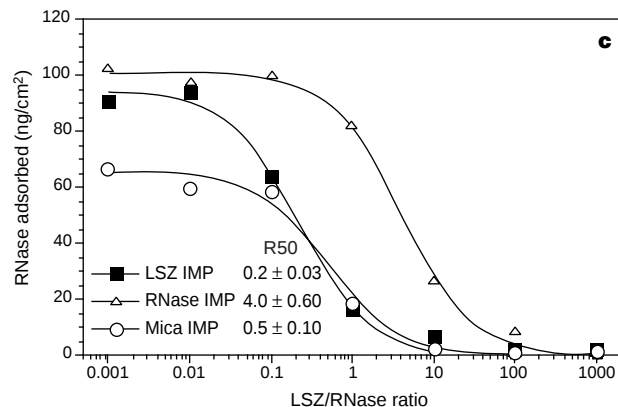
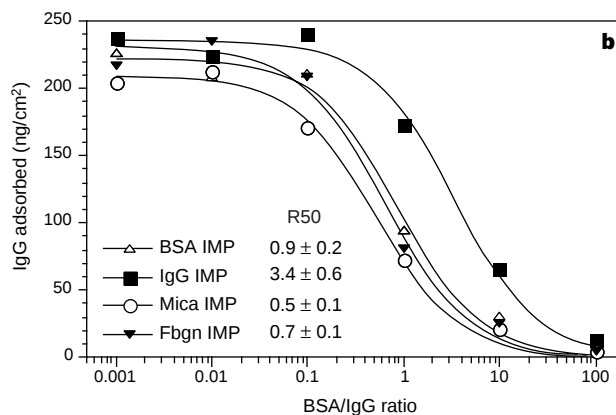
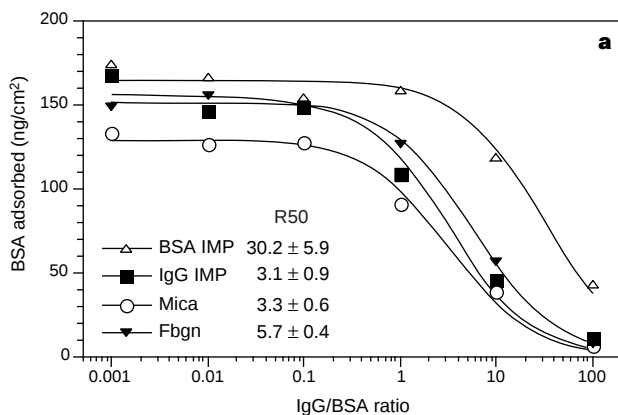


Figure 3 Competitive adsorption of a template protein versus a non-template protein was assayed by using a series of binary protein-mixture solutions. Template proteins were radiolabelled with [¹²⁵I]NaI by the ICI (iodine monochloride) technique²⁴. Protein adsorption was carried out in citrated PBS buffer, pH 7.4, for 2 h. Concentrations of competing, unlabelled protein solution were varied while the concentration of [¹²⁵I]-labelled protein was held constant. As more competing protein was added, less radiolabelled protein adsorbed onto the surface, leading to a sigmoidal adsorption

curve. **a–d**, Competitive adsorption of a binary protein mixture on imprint surfaces is shown for radiolabelled albumin (BSA) competing with unlabelled IgG (**a**); labelled IgG versus BSA (**b**); labelled ribonuclease A (RNase) versus lysozyme (LSZ) (**c**); and labelled LSZ versus RNase (**d**). Fbgn, fibrinogen. Data are averages of triplicate samples. Error bars are omitted for clarity. Kaleidagraph was used to fit curves to equations. The ratio required to cause a 50% reduction in the maximum labelled-protein adsorption (R_{50}) was determined for each imprint surface.

albumin (BSA; relative molecular mass 66,000 (M_r 66K)), immunoglobulin G (IgG; M_r 150K) and fibrinogen (M_r 330K), because of their important roles in mediating cellular responses to blood-contacting materials²³. Surface topography of the protein imprints was investigated by tapping-mode atomic force microscopy (AFM)¹⁶. Distinctive indentations related to the shape of template proteins were observed on the imprint surfaces, with the resolution being limited by the AFM tip dimensions. Figure 1b shows a fibrinogen imprint surface that has elongated trenches loosely resembling the shape of fibrinogen molecules. Heterogeneity in the image can be explained by a distribution of orientations and conformations of proteins that have been adsorbed onto mica. Statistical analysis of surface roughness and autocovariance indicated that fibrinogen imprints had the largest nanocavities, followed by IgG imprints and then BSA imprints, in agreement with the sizes of these proteins.

The chemical composition of the surface of protein imprints was investigated by electron spectroscopy for chemical analysis (ESCA) and time-of-flight secondary ion mass spectrometry (TOF-SIMS). The large decreases in the ESCA nitrogen signal and of TOF-SIMS

peaks—indicative of protein fragments—suggest that our basic solution eluted the embedded protein molecules out from the protein imprints (Fig. 2). A surface chemistry dominated by sugar species was seen for the protein imprints. Advancing contact angle (θ_a) measurements confirmed the hydrophilic nature of the protein-imprinted surfaces, which is consistent with the presence of sugar groups (θ_a of imprint surfaces was $35 \pm 9^\circ$).

¹²⁵I-labelled BSA, IgG and fibrinogen were adsorbed onto various protein imprints and controls. The amount of protein adsorbed was slightly higher on the protein imprints than on a mica-imprinted control surface (no template protein), suggesting that a larger surface area had been induced by the imprinted protein pits. However, no significant difference was detected between the amount of protein adsorbed on different protein imprints, which is consistent with nonspecific protein adsorption. Nonetheless, the amount of adsorbed protein retained by the surfaces following a detergent elution varied by imprint. A particular protein imprint generally retained more of its template protein than the imprint of another protein or an atomically flat (mica) imprint, implying that a template protein has a greater affinity for its imprint.

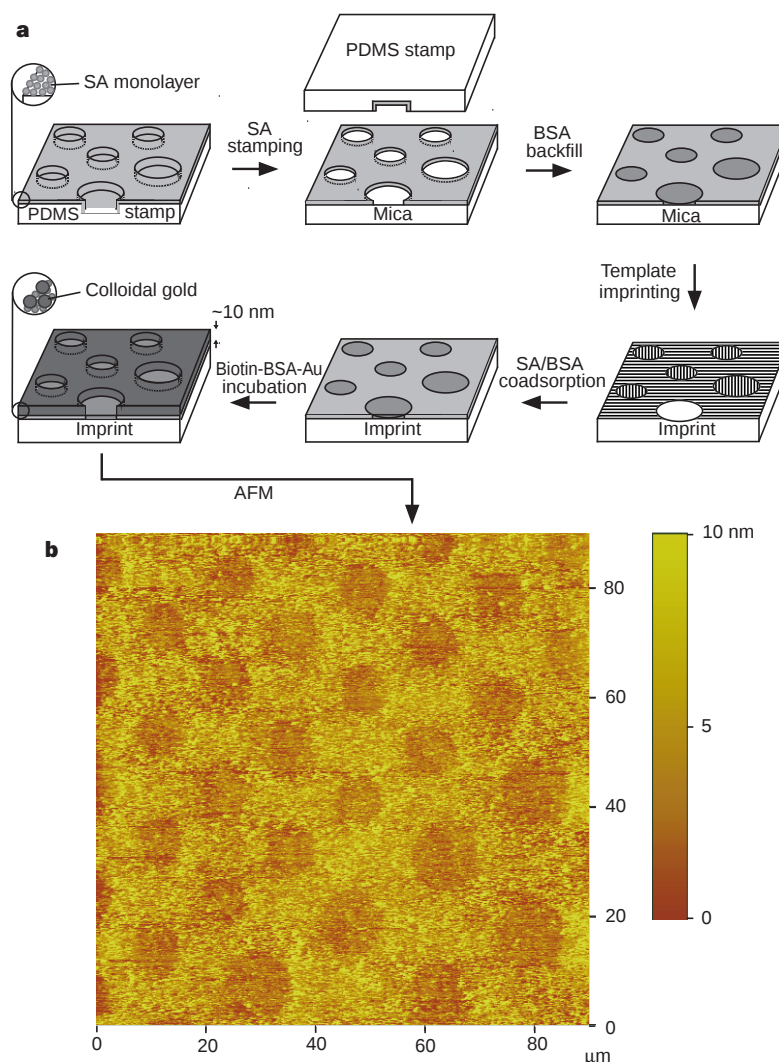


Figure 4 Visual demonstration of template recognition by imprints of protein mixtures patterned with microcontact printing^{29,30}. **a**, Streptavidin (SA) in PBS buffer (pH 7.4) was adsorbed onto a poly(dimethylsiloxane) (PDMS) stamp that has micrometre-scale circular indentations, followed by a buffer and water rinse and drying under a stream of nitrogen. The stamp was put into contact with a mica surface for ~5 s, transferring a monolayer of streptavidin to mica only in the contacted regions. The mica surface was then exposed to albumin (BSA) in PBS.

This adsorption blocks the uncovered surface area. A template imprint of this SA/BSA patterned surface was prepared as shown in Fig. 1. Binary protein adsorption onto the imprint was performed from a SA/BSA (1/1) solution in PBS for 4 h. The imprint was incubated 1 h with a solution of biotin-BSA labelled with 10-nm colloidal gold to reveal the surface distribution of adsorbed streptavidin with AFM (**b**).

Competitive adsorption of a binary protein mixture revealed a highly preferential adsorption of a template protein, BSA (Fig. 3a) or IgG (Fig. 3b), onto its own imprint. The higher selectivity of an imprint for its template protein is indicated by a significantly increased competition ratio required to cause a 50% reduction in the maximum adsorption (R_{50})²⁴ (5–10-fold increase for the BSA imprint and 4–7-fold for the IgG imprint). To investigate shape specificity in protein recognition further, we imprinted lysozyme (M_r 14.6K, $45 \times 30 \times 30 \text{ \AA}^3$) and ribonuclease A (RNase; M_r 13.7K, $38 \times 28 \times 22 \text{ \AA}^3$), many of whose physical chemical properties are similar, including their structural rigidity and isoelectric points²⁵. In RNase adsorption from the lysozyme/RNase mixture (Fig. 3c), there was a 20-fold increase in selectivity for RNase for the RNase imprint over the lysozyme imprint, indicating that template binding was selective. In adsorbing lysozyme from the RNase/lysozyme mixture (Fig. 3d), the lysozyme imprint showed a 26-fold enhanced selectivity for lysozyme over the RNase imprint, confirming the specificity of template recognition.

To visualize protein recognition directly, we imprinted BSA and streptavidin (M_r 65K) that had been spatially micropatterned on a mica surface (Fig. 4a). After competitive adsorption from a BSA/streptavidin mixture and incubation with biotin-conjugated 10-nm colloidal gold, the imprint surface was observed by AFM to have a pattern of colloidal particles, localized in the region where streptavidin molecules were preferentially adsorbed (Fig. 4b). This pattern was identical to that of the stamp used for patterning streptavidin, indicating that the streptavidin-imprinted surface region selectively adsorbed streptavidin molecules in the presence of BSA. The fact that streptavidin molecules adsorbed onto imprint surfaces can still bind biotin suggests that they sustain conformational integrity and therefore that their biological functionality is maintained. Thus, cavity-bound protein still retains active surface-binding domains.

The fact that protein recognition is evident only in competitive adsorption supports the idea that exchange is occurring between nonspecifically adsorbed non-template protein and the solution-phase template protein. Template imprints may be recognized by proteins through a multistep adsorption, with the specificity conferred by shape selectivity and hydrogen bonding. When protein molecules arrive at the surface, only a fraction of them stick or adsorb onto it^{26,27}. Compared with non-template proteins, a template protein entering its imprint will have a higher likelihood of being retained as a result of interlocking within a pit and subsequently binding strongly to it. In addition, adsorbed protein on a low-adsorptivity surface can exchange with dissolved protein in solution^{5,28}. Non-template protein that does not fit into a pit is more readily displaced than template protein²⁹, because the pit occupied by the template protein is no longer accessible to solution-phase protein. The hydrophilic, crosslinked sugars on protein imprints, in contrast to hydrophobic surfaces, allow for a lower protein-sticking probability and a higher protein exchangeability. Both of these processes lead to 'recognition of the fittest' through dynamic adsorption–exchange, which we believe is essential for protein recognition. □

Received 30 July 1998; accepted 23 February 1999.

- Vijayalakshmi, M. A. Pseudobiospecific ligand affinity chromatography. *Trends Biotechnol.* **7**, 71–76 (1989).
- Byfield, M. P. & Abuknesha, R. A. Biochemical aspects of biosensors. *Biosens. Bioelect.* **9**, 373–400 (1994).
- Ratner, B. D. The engineering of biomaterials exhibiting recognition and specificity. *J. Mol. Recogn.* **9**, 617–625 (1996).
- Ratner, B. D. New ideas in biomaterials science—a path to engineered biomaterials. *J. Biomed. Mat. Res.* **27**, 837–850 (1993).
- Brash, J. L. in *Biomaterials: Interfacial Phenomenon and Applications*, ACS Advances in Chemistry Series, 199 (eds Charm, C. P. & Szycher, M.) 3–24 (Technomic, Lancaster, PA, 1991).
- Mosbach, K. & Ramstrom, O. The emerging technique of molecular imprinting and its future impact on biotechnology. *Bio/technology* **14**, 163–170 (1996).
- Wulff, G. Molecular imprinting in cross-linked materials with the aid of molecular templates—a way towards artificial antibodies. *Angew. Chem. Int. Ed. Engl.* **34**, 1812–1832 (1995).
- Shea, K. J. Molecular imprinting of synthetic network polymers: the de novo synthesis of macro-molecular binding and catalytic sites. *Trends Polym. Sci.* **2**, 166–173 (1994).
- Steinke, J., Sherrington, D. & Dunkin, I. Imprinting of synthetic polymers using molecular templates. *Adv. Polym. Sci.* **123**, 80–125 (1995).
- Cram, D. J. The design of molecular hosts, guests, and their complexes. *Science* **240**, 760–767 (1988).

- Mallick, S. J. *et al.* Towards materials for the specific recognition and separation of proteins. *New J. Chem.* **18**, 299–304 (1994).
- Glad, M., Norrlov, O., Sellergren, B., Siegbahn, N. & Mosbach, K. Use of silane monomers for molecular imprinting and enzyme entrapment in polysiloxane-coated porous silica. *J. Chromatogr.* **347**, 11–23 (1985).
- Kempe, M., Glad, M. & Mosbach, K. An approach towards surface imprinting using the enzyme ribonuclease A. *J. Mol. Recogn.* **8**, 35–39 (1995).
- Venton, D. L. & Gudipati, E. Entrapment of enzyme using organo-functionalized polysiloxane copolymers. *Biochim. Biophys. Acta* **1250**, 117–125 (1995).
- Hjerten, S. *et al.* Gels mimicking antibodies in their selective recognition of proteins. *Chromatography* **44**, 227–234 (1997).
- Shi, H., Castner, D. G. & Ratner, B. D. High fidelity template imprinting of cells, proteins and colloidal gold particles. *Adv. Mater.* (submitted).
- Nock, S., Spudich, J. A. & Wagner, P. Reversible, site-specific immobilization of polyarginine-tagged fusion protein on mica surfaces. *FEBS Lett.* **414**, 233–238 (1997).
- Crowe, J. H., Carpenter, J. F., Crowe, L. M. & Anchordoguy, T. J. Are freezing and dehydration similar stress vectors? A comparison of modes of interaction of stabilizing solutes with biomolecules. *Cryobiology* **27**, 219–231 (1990).
- Lemieux, R. U. How water provides the impetus for molecular recognition in aqueous solution. *Acc. Chem. Res.* **29**, 373–380 (1996).
- Paleos, C. M. & Tsiourvas, D. Molecular recognition of organized assemblies via hydrogen bonding in aqueous media. *Adv. Mater.* **9**, 695–710 (1997).
- Ariga, K. & Kunitake, T. Molecular recognition at air–water and related interfaces: complementary hydrogen bonding and multisite interactions. *Acc. Chem. Res.* **31**, 371–378 (1998).
- Yu, C. & Mosbach, K. Molecular imprinting utilizing an amide functional group for hydrogen bonding leading to highly efficient polymers. *J. Org. Chem.* **62**, 4057–4064 (1997).
- Horbett, T. A. Principles underlying the role of adsorbed plasma proteins in blood interactions with foreign materials. *Cardiovasc. Pathol.* **2**(suppl.), 137–148 (1993).
- Horbett, T. A. in *Techniques in Biocompatibility Testing, II* (ed. Williams, D. F.) 183–214 (CRC, Boca Raton, FL, 1986).
- Shirahama, H., Lyklema, J. & Norde, W. Comparative protein adsorption in model systems. *J. Colloid Interface Sci.* **139**, 177–186 (1990).
- Weaver, D. R. & Pitt, W. G. Sticking coefficients of adsorbing proteins. *Biomaterials* **13**, 577–584 (1992).
- Norde, W. & Anusiem, A. C. I. Adsorption, desorption and re-adsorption of proteins on solid surfaces. *Colloids Surfaces* **66**, 73–80 (1992).
- Andrade, J. D. in *Surface and Interfacial Aspects of Biomedical Polymers: Proteins Adsorption, II* (ed. Andrade, J. D.) 1–80 (Plenum, New York, NY, 1985).
- Shi, H. & Ratner, B. D. Template recognition of protein imprinted polymer surfaces. *J. Biomed. Mat. Res.* (in the press).
- Bernard, A. *et al.* Printing patterns of proteins. *Langmuir* **14**, 2225–2229 (1998).
- James, C. D. *et al.* Patterned protein layers on solid substrates by thin stamp microcontact printing. *Langmuir* **14**, 741–744 (1998).

Acknowledgements. We thank T. A. Horbett, D. G. Castner, J. B. Lhoest and R. Luginbuhl for stimulating discussions, S. Golledge and I. Yu for assistance with ESCA and TOF-SIMS measurements, W. Ciridon for assistance in plasma deposition, K. Leach for help in contact angle measurement, and M. Sarikaya and H. K. Fong for access to the tapping mode AFM. This work was supported by NESAC/BIO(NIH) and UWEB(NSF). H.S. thanks the Center for Nanotechnology at the University of Washington for a fellowship.

Correspondence and requests for materials should be addressed to B.D.R. (e-mail: ratner@uweb.engr.washington.edu).

Influence of El Niño on the equatorial Pacific contribution to atmospheric CO₂ accumulation

Richard A. Feely*, Rik Wanninkhof†, Taro Takahashi‡ & Pieter Tans§

* Pacific Marine Environmental Laboratory, NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115-0070, USA

† Atlantic Oceanographic and Meteorological Laboratory, NOAA, 4301 Rickenbacker Causeway, Miami, Florida 33149, USA

‡ Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA

§ Climate Monitoring and Diagnostics Laboratory, 325 Broadway, Boulder, Colorado 80303, USA

The equatorial oceans are the dominant oceanic source of CO₂ to the atmosphere, annually amounting to a net flux of 0.7–1.5 Pg (10¹⁵ g) of carbon, up to 72% of which emanates from the equatorial Pacific Ocean^{1–3}. Limited observations indicate that the size of the equatorial Pacific source is significantly influenced by El Niño events^{4–10}, but the effect has not been well quantified. Here we report spring and autumn multiannual measurements of the partial pressure of CO₂ in the surface ocean and atmosphere in the equatorial Pacific region. During the 1991–94 El Niño period,

the derived net annual sea-to-air flux of CO₂ was 0.3 Pg C from autumn 1991 to autumn 1992, 0.6 Pg C in 1993, and 0.7 Pg C in 1994. These annual fluxes are 30–80% of that of 1996, a non-El Niño year. The total reduction of the regional sea-to-air CO₂ flux during the 1991–94 El Niño period is estimated to account for up to one-third of the atmospheric anomaly (the difference between the annual and long-term-average increases in global atmospheric CO₂ content) observed over the same period.

A comprehensive set of atmospheric and surface ocean CO₂ partial-pressure (p_{CO_2}) measurements and supporting hydrographic data were obtained during two field seasons in the boreal spring and autumn of 1992, the spring of 1993, the spring and autumn of 1994, and the spring and autumn of 1996. The spring cruises from 1992 to 1994 occurred during the well developed warm El Niño/Southern Oscillation (ENSO) event, and the autumn cruises took place during “cold-tongue” conditions¹¹.

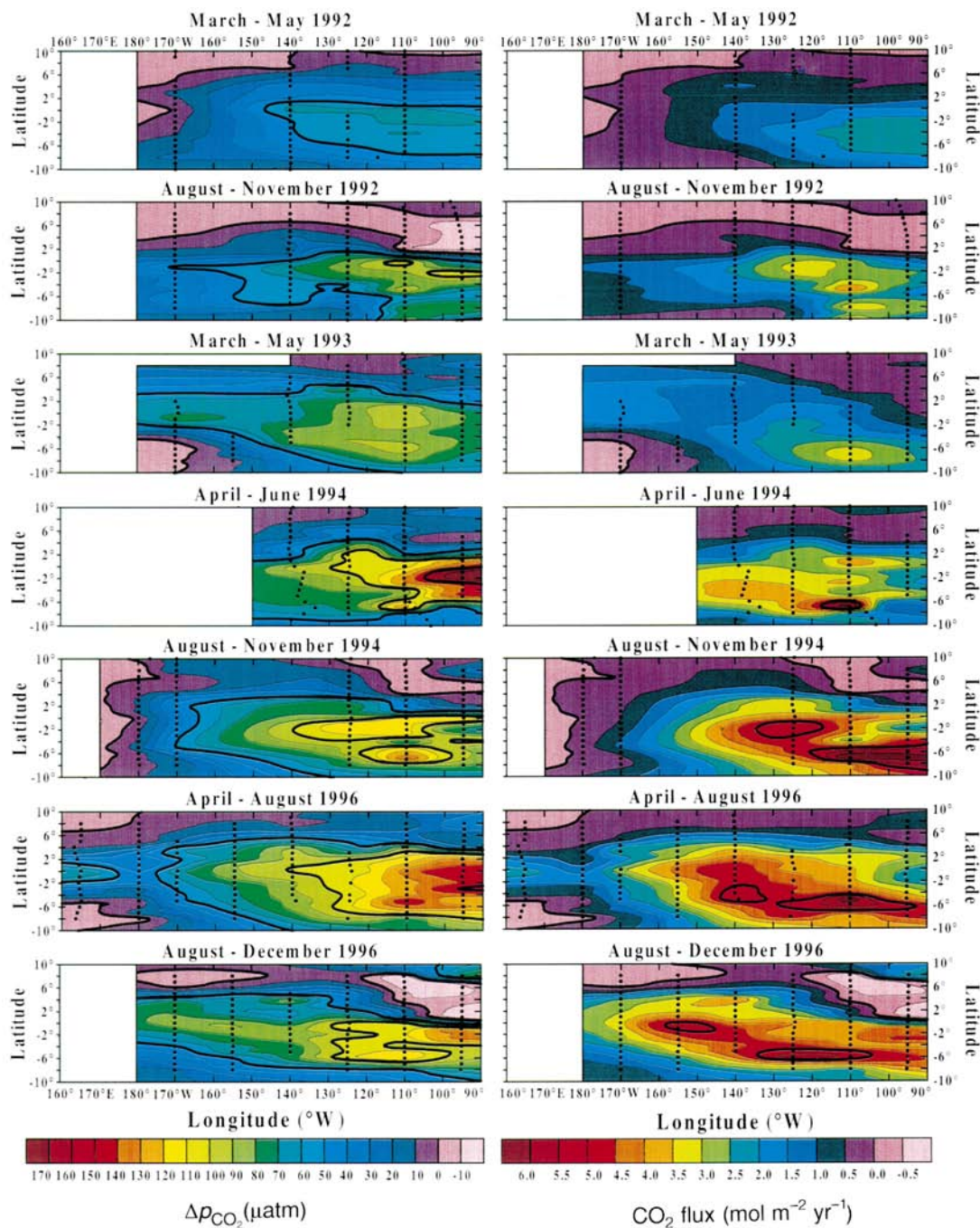


Figure 1 CO₂ disequilibria and sea-to-air fluxes in the equatorial Pacific. The distributions of Δp_{CO_2} (see text) in μatm (left) and CO₂ flux in $\text{mol m}^{-2} \text{yr}^{-1}$ (right) in the central and eastern equatorial Pacific are for two field seasons in the boreal spring and autumn of 1992; the spring of 1993; the spring and autumn of 1994; and spring/summer and autumn/winter of 1996. The lower Δp_{CO_2} values during the 1992–94 ENSO event, combined with lower wind speeds, resulted in a significant decrease in the CO₂ flux compared with the 1996 non-El Niño year. ‘Underway’ atmospheric and surface ocean p_{CO_2} measurements and supporting hydrographic data were obtained during these cruises on board the NOAA Ships

Discoverer, *Malcolm Baldrige* and *Ka'imimoana*. The ‘underway’ measurements of CO₂ mixing ratios in dry air were made with a Li-Cor (Model 6251) non-dispersive infrared analyser linked to the shipboard equilibrator following the methods described in ref. 29. The system runs on an hourly cycle during which standards, a headspace sample from the equilibrator, and an ambient air sample are analysed. Standards were obtained from NOAA's Climate Monitoring Diagnostics Laboratory (CMDL) at Boulder. All reference standards undergo a pre- and post-cruise calibration at CMDL against standards certified by the World Meteorological Organization (WMO).

The eastern equatorial Pacific Ocean is characterized by high seawater p_{CO_2} and high nutrient concentrations driven by upwelling of the Equatorial Undercurrent. This water mass mixes with a combination of the westward-flowing South Equatorial Current and upwelled subtropical water, which is transported poleward by Ekman divergence along the Equator^{11–13}. As a consequence of these processes, this region is an important site for the exchange of carbon between the ocean interior and the atmosphere. Although previous field studies of seawater p_{CO_2} in this region have emphasized the high degree of interannual variability caused by the decline of upwelling during ENSO events^{4–10}, the observations were limited in their spatial and temporal coverage. Our 1996 data set clearly indicates higher Δp_{CO_2} (seawater p_{CO_2} – air p_{CO_2}) values than 1992 and also that the region covered by positive Δp_{CO_2} extends farther in both the north–south and east–west directions (Fig. 1). Whereas values close to equilibrium are observed north and south of the Equator in the western part of the study region, the region between 7° N and 6° S has a large positive Δp_{CO_2} centred on the Equator at 165° E in 1996. In contrast, the data for spring 1992 indicate near-equilibrium Δp_{CO_2} values in the region of 170° W. Both the seasonal and interannual distributions of seawater p_{CO_2} ($p_{\text{CO}_2\text{w}}$) in the eastern Pacific show decreasing trends from east to west, while the largest interannual variability is observed in the eastern Pacific along 95° W and 110° W.

Near the Equator there is a general inverse relationship between the $p_{\text{CO}_2\text{w}}$ maximum and temperature minimum (Fig. 2). This is caused by enhanced upwelling and advection of CO_2 -enriched waters during non-El-Niño periods^{4,5}. Sea surface p_{CO_2} is primarily affected by subsurface upwelling processes, which are controlled by local and remotely forced wind events^{11–13}. When the winds are

strongest during the cold cycle of ENSO deep upwelling occurs and $p_{\text{CO}_2\text{w}}$ values are at a maximum⁴.

The net flux of CO_2 across the sea–air interface is the product of the gas transfer velocity, k , the solubility of CO_2 , s , and the p_{CO_2} difference between the ocean and atmosphere ($p_{\text{CO}_2\text{w}} - p_{\text{CO}_2\text{a}}$), according to the following equation:

$$F = ks(p_{\text{CO}_2\text{w}} - p_{\text{CO}_2\text{a}}) \quad (1)$$

In order to provide a simple relationship for the wind-speed dependence that provided non-zero gas fluxes at low wind speeds, as well as considering the long-term variability of the winds over the open ocean, one of us (R.W.) developed¹⁴ a quadratic equation for the wind speed dependence that was fitted through the mean gas exchange based on the oceanic bomb ^{14}C data in which

$$k = 0.39w_{10}^2(\text{Sc}/660)^{-0.5} \quad (2)$$

where Sc is the Schmidt number for CO_2 , and w_{10} is the wind speed at 10 m height in (m s^{-1}). We used this relationship for the CO_2 flux calculations. One-month averaged re-analysis wind speeds were acquired from the European Centre for Medium Range Weather Forecasts.

Spring flux data show large interannual effects of El Niño on CO_2 exchange in the equatorial Pacific (Fig. 1). From the peak El Niño period during spring of 1992 to the normal conditions in spring of 1996, the average CO_2 fluxes from 10° S to 10° N and 90° W to 140° W increased from approximately 1.2 to 3.1 $\text{mol m}^{-2} \text{yr}^{-1}$. The increased CO_2 fluxes in 1994 and 1996 are due to increased surface-water p_{CO_2} values and higher winds, resulting in stronger upwelling of CO_2 -enriched water from depths of about 70–100 m within the upper portion of the Equatorial Undercurrent. Figure 1 also shows a springtime north–south asymmetry, with the highest fluxes occurring in the southeastern portion of the study area due to the stronger southeasterly winds which enhance localized upwelling. These results indicate that the CO_2 flux to the atmosphere over the region south of the Equator is nearly twice as large as that north of it during spring. The 1992 autumn fluxes were also significantly lower in the eastern Pacific as compared with the autumn of 1996. Thus, the 1991–94 ENSO event was a period when the net sea-to-air flux was significantly lower than the non-El-Niño period in 1996.

The CO_2 fluxes during the 1992 period were very similar to those observed during the peak of the 1986–87 ENSO event in the equatorial Pacific⁸. But the significant feature of the 1991–94 ENSO event was that it extended from the autumn of 1991 to the beginning of 1994. Table 1 gives a summary of estimates of annual CO_2 fluxes from the equatorial Pacific for the period 1992–1994 and 1996 over the region from 10° S to 10° N and from 80° W to 135° E. For the data from 1992 to 1994, we used the data from this study for the central and eastern Pacific. For the western Pacific we used the $p_{\text{CO}_2\text{w}}$ gradients given in Inoue *et al.*¹⁵ The results show that for the 1-year period from the autumn of 1991 until the autumn of 1992, approximately 0.3 Pg C was released to the atmosphere. During the 12-month period of 1993 approximately 0.6 Pg C was released to the atmosphere, in 1994 0.7 Pg C, and in 1996 0.9 Pg C. The total decrease of CO_2 outgassing from the equatorial oceans during the prolonged 1991–94 El Niño period (that is, 0.8–1.2 Pg C) was larger than during the severe El Niño of 1982–83.

Recently, atmospheric chemists have reported a significant decrease in the annual atmospheric growth rate of CO_2 between 1991 and 1993, from a mean value of 1.5 parts per million per year (p.p.m. yr^{-1}) throughout the 1980s to ~ 0.7 p.p.m. yr^{-1} in 1992^{16–19}. This was followed by normal values during the 1994–1996 period (average growth rate, 1.65 p.p.m. yr^{-1}). Although all groups have indicated that the slowdown of the atmospheric CO_2 growth rate must be the result of an enhanced uptake of CO_2 in the oceans and the terrestrial biosphere, there are considerable differences about the role of the oceans in this process. Francey *et al.*¹⁶ used atmospheric p_{CO_2} and $^{13}\text{C}/^{12}\text{C}$ isotope ratio data to estimate temporal variations

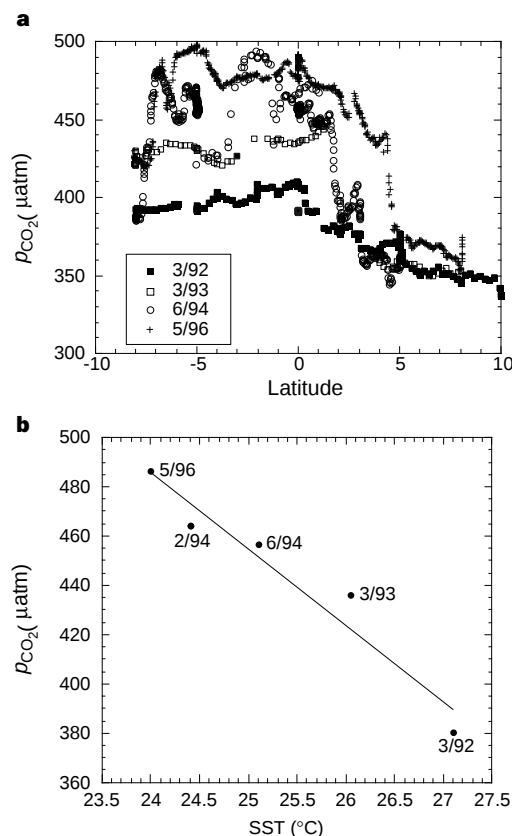


Figure 2 Time series of $p_{\text{CO}_2\text{w}}$. The data are from cruises during the boreal spring between 1992–1994 and 1996 (SST, sea surface temperature). The month of observations is listed in the key. **a**, Time series along 110° W. **b**, The $p_{\text{CO}_2\text{w}}$ versus the temperature at the temperature minimum along 110° W.

Table 1 Estimates of the equatorial Pacific sea-air CO₂ flux

Year	Annual equatorial Pacific flux (Pg C)	Annual equatorial Pacific flux anomaly* (Pg C)	Annual atmospheric flux anomaly† (Pg C)	Per cent of atmospheric anomaly due to the equatorial Pacific ENSO
1992	0.3 ± 0.2	-0.6 ± 0.4	-1.69	36
1993	0.6 ± 0.4	-0.3 ± 0.3	-1.84	16
1994	0.7 ± 0.5	-0.2 ± 0.3	0.09	NA
1996	0.9 ± 0.6	0.0	-0.03	NA

NA, Not applied.

*Equatorial Pacific flux – 1996 equatorial Pacific flux. From an analysis of seasonal and interannual datasets the uncertainties are composed of the seasonal variations of $p_{\text{CO}_2\text{w}}$ (28%), wind speed variability (28%), and the wind-speed dependence of gas exchange (44%). The region of coverage for the equatorial Pacific flux estimates is from 10° S to 10° N and from 80° W to 135° E. †Difference between measured atmospheric annual CO₂ increase and a long-term average annual increase of 3.18 Pg C (ref. 19).

of terrestrial and oceanic sinks for CO₂ from 1982 through 1992. Compared to earlier years, their 1992 global oceanic sink is about 2 Pg C yr⁻¹ larger than average. Keeling *et al.*¹⁸, using a similar global CO₂ and isotopic mass balance based on different data sets, found the ocean sink to be 0.6 Pg C yr⁻¹ larger than average in 1992 and 0.9 Pg C yr⁻¹ smaller in 1993. Our data for the equatorial Pacific suggests that this region plays a significant role in the decreased sea-to-air flux of CO₂ during an ENSO event. The decreased amount of CO₂ released to the atmosphere in the equatorial Pacific during the 1991–94 El Niño period (that is, 0.8–1.2 Pg C) was approximately 16–36% of the atmospheric anomaly observed over the same period (Table 1). Other oceanic regions could have contributed to the enhanced atmospheric anomaly¹⁹. However, the results of this study represent the largest documented interannual change in the oceanic CO₂ source flux observed to date.

Model predictions of Δp_{CO_2} in the equatorial Pacific have suggested a decrease over time because the $p_{\text{CO}_2\text{a}}$ is increasing much faster from anthropogenic inputs than is the p_{CO_2} of the upwelled water²⁰. This decrease would seem to be reasonable if the main source of the upwelled water is the deep core of the Equatorial Undercurrent which originates in the deep (>300 m) waters of the western South Pacific^{21,22}. The time lag for ventilation of anthropogenic CO₂ into the formation waters and its transport to the equatorial Pacific would cause a slower rate of CO₂ increase in the upwelled water. However, other models and tracer observations of equatorial upwelling indicate that a significant fraction of the upwelled water is either recycled surface water or entrained water from shallow subtropics north and south of the Equator^{23,24}. Under these conditions, surface ocean water p_{CO_2} should have been nearly equilibrated with atmospheric CO₂ and increased at nearly the same rate as atmospheric CO₂. To validate these models, we evaluated seven data sets from the eastern equatorial Pacific collected between 140° W

and 160° W in the spring during non-El-Niño years. All data were converted to *in situ* p_{CO_2} using the equations outlined in Murphy *et al.*²⁵ For each cruise, we constructed a plot of $p_{\text{CO}_2\text{w}}$ versus sea surface temperature for the core of the upwelled water in the region of the temperature minimum near the Equator, nominally between 0.5° N and 0.5° S. We then made a linear extrapolation of the data to a constant temperature of 25°C to remove the effects of solubility changes due to differences in seawater temperature. The results are plotted as a function of time in Fig. 3. The atmospheric data show a rate of increase of about 1.25 $\mu\text{atm yr}^{-1}$ between 1961 and 1996, with a standard error of $\pm 0.06 \mu\text{atm yr}^{-1}$. The oceanic data also show a similar rate of increase ($\sim 1.27 \mu\text{atm yr}^{-1}$), although the uncertainty is higher (standard error $\pm 0.18 \mu\text{atm yr}^{-1}$). A similar constancy of Δp_{CO_2} near 150° W was observed by Goyet and Peltzer⁹ for two cruises in the autumn of 1979 and 1991. As the total range of the $p_{\text{CO}_2\text{w}}$ difference ($\sim 50 \mu\text{atm}$) over the 35-year interval in Fig. 3 exceeds the normal local short-term variations of temperature-corrected $p_{\text{CO}_2\text{w}}$ in this region²⁶, we conclude that the long-term $p_{\text{CO}_2\text{w}}$ increases are due to anthropogenic inputs. For this data set, it appears that Δp_{CO_2} normalized to 25°C has remained constant at about $100 \pm 3 \mu\text{atm}$ over the period of the observations, providing further support for the shallow recycling mechanism. It appears that only a portion of the upwelled water at the surface originates from the core of the Equatorial Undercurrent, but that this is sufficient to supply a significant fraction of the nutrient concentrations observed at the surface.

The oceanic increase of $p_{\text{CO}_2\text{w}}$ in the equatorial Pacific agrees reasonably well with other time-series measurements for the subtropical waters. From time-series measurements of $p_{\text{CO}_2\text{w}}$ in the western subtropical Pacific between 3° N and 30° N from 1984 to 1993, Inoue *et al.*²⁷ obtained an average growth rate of $1.2 \pm 0.9 \mu\text{atm yr}^{-1}$. Similarly, the increase with time of oceanic CO₂ of approximately $1.5 \mu\text{atm yr}^{-1}$ has been observed at the HOT time-series station off Hawaii²⁸. These results support our belief that the entrained subtropical water which is injected into the upwelled water at the Equator has been recently exposed to, and equilibrated with, the atmosphere. □

Received 5 November 1998; accepted 10 February 1999.

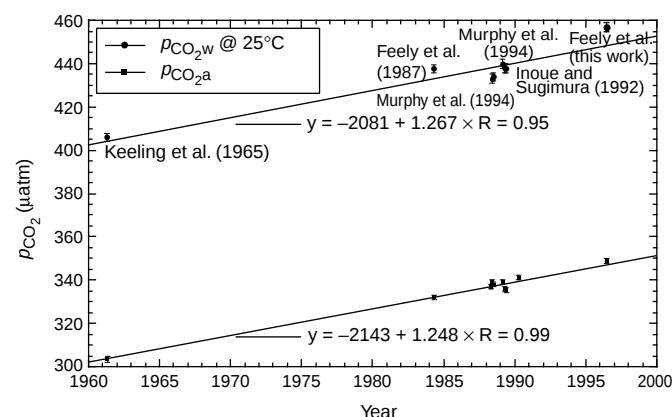


Figure 3 Long time series of seawater and atmospheric p_{CO_2} . The region of coverage is the central equatorial Pacific between 140° W and 160° W. The seawater values have been corrected to a constant temperature of 25°C to remove the effect of temperature on CO₂ solubility. Data sources: 1961, Keeling *et al.*³⁰; 1984, Feely *et al.*⁷; 1988, Murphy *et al.*²⁵; 1989, Murphy *et al.*²⁵; 1989, Inoue and Sugimura⁸; 1996, Feely *et al.* (this work).

- Houghton, J. *et al.* (eds) *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC 1992 Emission Scenarios* (Cambridge Univ. Press, 1994).
- Tans, P. P., Fung, I. Y. & Takahashi, T. Observational constraints on the global atmospheric CO₂ budget. *Science* **247**, 1431–1438 (1990).
- Takahashi, T. *et al.* Global air–sea flux of CO₂: An estimate based on measurements of sea–air pCO₂ difference. *Proc. Natl Acad. Sci. USA* **94**, 8292–8299 (1997).
- Feely, R. A. *et al.* CO₂ distributions in the equatorial Pacific during the 1991–92 ENSO event. *Deep-Sea Res.* **42**(2–3), 365–386 (1995).
- Wanninkhof, R., Feely, R. A., Chen, H., Cosca, C. & Murphy, P. P. Surface water fCO₂ values in the eastern equatorial Pacific during the 1992–93 El Niño. *J. Geophys. Res.* **101**, 16333–16343 (1996).
- Murray, J. W., Barber, R. T., Roman, M., Bacon, M. & Feely, R. A. Physical and biological controls on carbon cycling in the equatorial Pacific. *Science* **266**, 58–65 (1994).
- Feely, R. A. *et al.* Distribution of chemical tracers in the eastern equatorial Pacific during and after the 1982–1983 El Niño/Southern Oscillation event. *J. Geophys. Res.* **92**, 6545–6558 (1987).
- Inoue, H. & Sugimura, Y. Variations and distributions of CO₂ in and over the equatorial Pacific during the period from 1986/88 El Niño event to the 1988/89 La Niña event. *Tellus B* **44**, 1–22 (1992).
- Goyet, C. & Peltzer, E. T. Comparison of the August–September 1991 and 1979 surface partial pressure of CO₂ distribution in the Equatorial Pacific Ocean near 150° W. *Mar. Chem.* **45**, 257–266 (1994).
- Archer, D. E. *et al.* Seasonal and interannual variability of sea-surface carbon and nutrient concentration in the equatorial Pacific Ocean. *Deep-Sea Res.* **43**(4–6), 779–808 (1996).
- McPhaden, M. J. TOGA-TAO and the 1991–92 ENSO event. *Oceanography* **6**(2), 36–44 (1993).

12. Harrison, D. E. Vertical velocity in the central tropical Pacific—a circulation model perspective for JGOFS. *Deep-Sea Res. II* **43**(4–6), 687–705 (1996).
13. Toggweiler, J. R. & Carson, S. in *Upwelling the Ocean: Modern Processes and Ancient Records* (eds Summerhayes, C. et al.) 337–360 (Wiley, Chichester, 1995).
14. Wanninkhof, R. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.* **97**, 7373–7382 (1992).
15. Inoue, H. Y., Ishii, M., Matsueda, H. & Aoyama, M. Changes in longitudinal distributions of partial pressure of CO₂ (pCO₂) in the central and western equatorial Pacific, west of 160°W. *Geophys. Res. Lett.* **23**(14), 1781–1784 (1996).
16. Francey, R. J. et al. Changes in oceanic and terrestrial carbon uptake since 1982. *Nature* **373**, 326–330 (1995).
17. Ciais, P. et al. Partitioning of ocean and land uptake of CO₂ as inferred by CO₂ measurements from the NOAA/CMDL global air sampling network. *J. Geophys. Res.* **100**, 5051–5070 (1995).
18. Keeling, C. D., Whorf, T. P., Wahlen, M. & Vander Plicht, J. Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature* **375**, 666–670 (1995).
19. Ciais, P., Tans, P. P., Trolter, M., White, J. W. & Francey, R. J. A large northern hemisphere terrestrial CO₂ sink indicated by the ¹³C/¹²C ratio of atmospheric CO₂. *Science* **269**, 1098–1102 (1995).
20. Maier-Reimer, E. & Hasselmann, K. Transport and storage of CO₂ in the ocean—An inorganic ocean-circulation carbon cycle model. *Clim. Dyn.* **2**, 63–90 (1987).
21. Tsuchiya, M., Lukas, R., Fine, R. A., Firing, E. & Lindstrom, E. Source waters of the Pacific Equatorial Undercurrent. *Prog. Oceanogr.* **23**, 101–147 (1990).
22. Toggweiler, J. R., Dixon, K. & Broecker, W. S. The Peru upwelling and the ventilation of the South Pacific thermocline. *J. Geophys. Res.* **96**, 20467–20497 (1991).
23. Quay, P. D. & Stuiver, M. Broecker upwelling rates for the equatorial Pacific Ocean derived from bomb ¹⁴C distribution. *J. Mar. Res.* **41**, 769–792 (1983).
24. Gu, D. & Philander, S. G. H. Interdecadal climate fluctuations that depend on exchanges between the tropics and extratropics. *Science* **275**, 805–807 (1997).
25. Murphy, P. P., Kelly, K. C., Feely, R. A. & Gammon, R. H. *Carbon Dioxide Concentrations in Surface Water and the Atmosphere: PMEL Cruises 1986–1989* (Tech. Memo. ERL PMEL-101 (DE95-009659), NOAA, Seattle, 1994).
26. Feely, R. A., Wanninkhof, R., Goyet, C., Archer, D. E. & Takahashi, T. Variability of CO₂ distributions and sea–air fluxes in the central and eastern equatorial Pacific during the 1991–94 El Niño. *Deep-Sea Res. II* **44**, 1851–1867 (1997).
27. Inoue, H. Y. et al. Long-term trend of the partial pressure of carbon dioxide (pCO₂) in surface waters of the western North Pacific, 1984–1993. *Tellus B* **47**, 391–413 (1995).
28. Winn, C. D., Li, Y.-H., Mackenzie, F. T. & Karl, D. M. Rising surface dissolved inorganic carbon at the Hawaii Ocean Time-series site. *Mar. Chem.* **60**(1–2), 33–48 (1998).
29. Wanninkhof, R. & Thoning, K. Measurement of CO₂ in surface water using continuous and discrete sampling methods. *Mar. Chem.* **44**, 189–205 (1993).
30. Keeling, C. D., Rakestraw, N. W. & Waterman, L. S. Carbon dioxide in the surface waters of the Pacific Ocean, 1—Measurements of the distribution. *J. Geophys. Res.* **70**, 6087–6097 (1965).

Acknowledgements. The assistance of the officers and crew of NOAA ships *Discoverer*, *Malcolm Baldrig*, and *Ka'imimoana* is acknowledged. We thank D.E. Harrison for wind data, and C. Cosca, M. Steckley and D. Ho for the pCO₂ measurements and data reduction. We also thank the NOAA TAO Project Office (director M. McPhaden) for support. This work was supported by the NOAA Climate and Global Change Program as part of the Ocean–Atmosphere Carbon Exchange Study.

Correspondence and requests for materials should be addressed to R.A.F. (e-mail: feely@pmel.noaa.gov).

Tectonic processes in Papua New Guinea and past productivity in the eastern equatorial Pacific Ocean

Mark. L. Wells^{*,‡}, Geoffrey K. Vallis[†] & Eli A. Silver[‡]

^{*}School of Marine Sciences, University of Maine, Orono, Maine 04469, USA

[‡]Institute of Marine Sciences, University of California, Santa Cruz, California 95064, USA

[†]GFDL, Princeton University, Princeton, New Jersey 08544, USA

Phytoplankton growth in the eastern equatorial Pacific Ocean today accounts for about half of the 'new' production—the fraction of primary production fuelled by externally supplied nutrients—in the global ocean. The recent demonstration that an inadequate supply of iron limits primary production in this region¹ supports earlier speculation that, in the past, fluctuations in the atmospheric deposition of iron-bearing dust may have driven large changes in productivity². But we argue here that only small (~2 nM) increases in the iron concentration in source waters of the upwelling Equatorial Undercurrent are needed to fuel intense diatom production across the entire eastern equatorial Pacific Ocean. Episodic increases in iron concentrations of this magnitude or larger were probably frequent in the past because a large component of the undercurrent originates in the convergent island-arc region of Papua New Guinea, which has experienced intensive volcanic, erosional and seismic activity over the past 16

million years. Cycles of plankton productivity recorded in eastern equatorial Pacific sediments may therefore reflect the influence of tectonic processes in the Papua New Guinea region superimposed on the effects of global climate forcing.

High-resolution sediment records in the eastern equatorial Pacific show strong cycling in silica:carbonate ratios resulting largely from the variable accumulation of diatom (silica) and coccolithophore (carbonate) tests³. These alternating conditions today reflect regimes having higher and lower productivity, respectively. Superimposed on these cycles are repeated episodes of extremely intense diatom production, manifested as laminated diatom deposits⁴. Increased upwelling of macronutrients driven by intensified global winds was thought to have been responsible for higher-productivity periods³, but a recent mesoscale enrichment study (IronEx II) showed¹ that diatom growth is limited by iron rather than inadequate supplies of nitrogen, phosphorus or silicon. The questions are now what magnitude of Fe inputs are needed to generate diatom bloom events and what was the frequency of these excursions?

Gauging the Fe perturbations required to stimulate diatom production reduces to two issues: (1) what is the lowest Fe concentration whereby uptake still permits maximum cellular growth rates, and (2) what Fe flux to surface waters is then needed to sustain high diatom productivity once this concentration threshold is surpassed? The distinction between concentration and flux is important, because it constrains the magnitude of events needed to generate intense diatom blooms. For example, maintaining relatively high (≥1 nM) Fe concentrations would require massive and persistent dust deposition because Fe is rapidly lost from surface waters⁵. In this case, significant global climate change must precede changes in diatom production. Conversely, if only low Fe concentrations are needed then flux requirements are substantially smaller, so that large shifts in productivity might

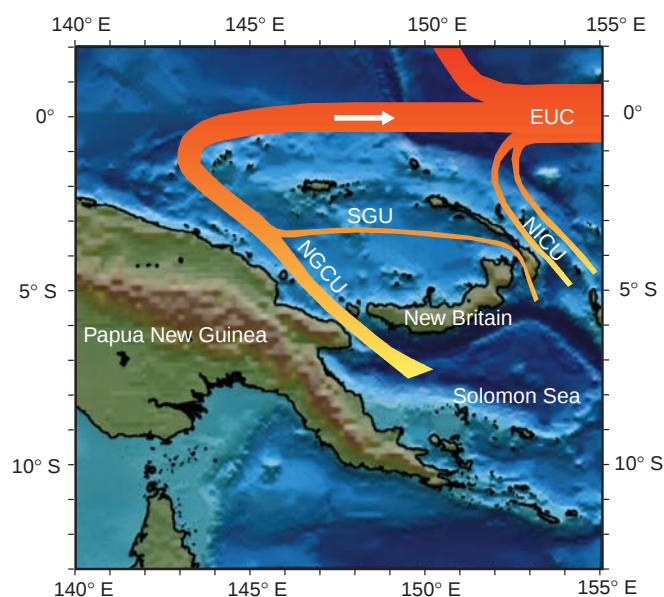


Figure 1 Source waters of the Equatorial Undercurrent (EUC). The New Guinea Coastal Undercurrent (NGCU) transports water at 200 m depth from the Solomon Sea, through the narrow Vitiaz Strait, and along the northern shelf margin of Papua New Guinea where it is joined by the St Georges Undercurrent (SGU)^{10,11}. The combined undercurrent turns north and then eastwards as the equatorial undercurrent. Additional undercurrents flowing north of New Ireland (the New Ireland Coastal Undercurrents, NICU), join the EUC, increasing its total flow. Subsurface waters north of the Equator also contribute to undercurrent flow; present estimates of the magnitude of this entrainment vary from small to roughly equal that of the southern inflows^{10,11}. Surface currents have been omitted for clarity.

Table 1 Iron upwelling fluxes needed to support diatom growth

Cellular Fe:carbon* ($\mu\text{mol Fe}:\text{mol C}$)	Fe upwelling flux† ($\text{nmol Fe m}^{-2} \text{d}^{-1}$)	Initial dissolved Fe‡ (nM)
2.4	360 (720)	2.1 (4.2)
4.0	600 (1,200)	3.4 (6.8)
6.0	900 (1,800)	5.2 (10.4)

* The range of cellular Fe quotas for oceanic phytoplankton determined in laboratory experiments⁷ and natural population cultures from equatorial Pacific surface waters¹⁵.

† Fe fluxes needed to support a calculated upwelling-nitrate-sustained new production rate of $150 \text{ nmol C m}^{-2} \text{d}^{-1}$ at 90°W (off the Galapagos) as a function of diatom Fe quotas. The prerequisite Fe flux would decrease westwards along the Equator due to lower nitrate upwelling fluxes. Values in parenthesis are for maximum, non-sustainable bloom production ($300 \text{ nmol C m}^{-2} \text{d}^{-1}$), such as occurred during IronEx II (that is, where nitrate concentration is drawn quickly to zero).

‡ The dissolved Fe concentrations in the PNG coastal undercurrent needed to support the required Fe upwelling flux at 90°W (near the eastern margin of the equatorial high-nitrogen, low-chlorophyll region).

possibly precede or occur independently of global climate change.

The threshold concentration where Fe uptake becomes diffusion-limited has been explored in models⁶ and contrived, well-defined laboratory culture experiments⁷. But predicting diffusion-limited conditions in ocean waters is difficult because we do not know which soluble chemical Fe species are directly accessed by phytoplankton⁸. Nonetheless, ambient soluble Fe concentrations (of species with molecular mass $<1 \text{ kDa}$) only doubled to 40 pM Fe during IronEx II despite nanomolar Fe additions; the bulk of added Fe was sequestered into colloidal forms not directly accessible to the cell (M.L.W., manuscript in preparation). Thus it appears that massive dust deposition is not a prerequisite for dramatically altering productivity in the eastern equatorial Pacific.

The bulk of Fe input to equatorial surface waters occurs by advective upwelling from the Equatorial Undercurrent (EUC) rather than atmospheric deposition⁹. The possibility that undercurrent Fe inputs have varied during the past has not been explored. The EUC originates near Papua New Guinea (PNG) as a coalescence of at least four undercurrents, important components of which are the New Guinea Coastal Undercurrent (NGCU) and the St Georges Undercurrent (SGU) (Fig. 1). Both draw water from the Solomon Sea along the north coast of PNG before turning towards the Equator at $\sim 145^\circ \text{E}$ ^{10,11}. This core undercurrent is then joined by the New Ireland Coastal Undercurrents (NICU in Fig. 1) plus subsurface contributions from northern waters to form the EUC. The $\sim 500\text{-km}$ -wide, 100-m -thick EUC then travels $\sim 13,000 \text{ km}$ constrained between 2°N and 2°S at speeds of $0.6\text{--}1.7 \text{ m s}^{-1}$. It begins its equatorial transit at $\sim 200 \text{ m}$ depth, but shoals gently upwards to surface south of the Galapagos Islands in the east. The New Guinea Coastal Undercurrent, the St Georges Undercurrent and the New Ireland Undercurrents are likely to strongly influence the dissolved Fe content of the EUC due to their very close proximity to land.

The loss of dissolved Fe from the ribbon-like EUC will largely result from advective upwelling into wind-mixed surface waters; turbulent diffusional losses will be small even with large ($\sim 10 \text{ cm}^2 \text{ s}^{-1}$) values of the vertical turbulent diffusivity coefficient. Using first-order approximations to model this loss (see Methods) suggests that Fe concentration in the upper EUC will fall exponentially during its passage across the Pacific Ocean. Lateral entrainment of water into the undercurrent¹² will probably further dilute the Fe concentration, perhaps by an additional factor of two. Scavenging losses to sinking particles will be small in relation to advective upwelling, given the comparatively low particle flux and high undercurrent speeds.

How well does this approach work? In this simplified model, Fe concentrations in the lower portion of the EUC will change little from those in the source waters while values in the upper segment decrease eastwards. Choosing a concentration of 0.35 nM Fe at 145°E yields a predicted value of $\sim 0.1 \text{ nM}$ Fe in the upper portion of the EUC at 140°W , roughly halfway across the Pacific. These two values are nearly identical to those measured in the upper and lower portions of the EUC in this region⁹. Although no Fe data have been

reported for the New Guinea Coastal Undercurrent, the St Georges Undercurrent or the New Ireland Coastal Undercurrents, the 0.35 nM value is remarkably close to that calculated (0.34 nM Fe) from reported oxygen values¹⁰ and a general $\text{O}_2\text{:Fe}$ relationship in the oceans¹³. Even with dilution of the EUC by entrainment during passage across the Pacific, this simplistic model provides a reasonable first approximation for calculating Fe upwelling fluxes across the eastern equatorial Pacific as a function of Fe concentrations in the coastal undercurrent.

We can judge the sensitivity of equatorial production to changes in dissolved Fe concentrations in EUC source waters by comparing model estimates of Fe upwelling fluxes with the Fe requirements for diatom growth. At 140°W , the upwelling flux of nitrate corresponds to a potential nitrate-sustained carbon production rate of $103 \text{ mmol C m}^{-2} \text{d}^{-1}$ (ref. 9). A gradual 50% increase eastwards in the fugacity of CO_2 in surface water¹⁴ leads to corresponding nitrate-sustained production rates of $\sim 150 \text{ mmol C m}^{-2} \text{d}^{-1}$ on the eastern margin of the equatorial Pacific. Seasonal variations in these estimates during a normal year will be comparatively small.

The calculated Fe flux needed to ensure that the upwelling nitrate fuels rapid diatom growth depends upon their cellular Fe quotas. Using the range of $2.4\text{--}6.2 \mu\text{mol Fe}:\text{mol C}$ measured in laboratory cultures⁷ and natural populations¹⁵ of oceanic diatoms, upwelling Fe fluxes of $360\text{--}900 \text{ nmol m}^{-2} \text{d}^{-1}$ would be required to fully nourish diatom production (Table 1). The onset of rapid growth would occur 2–5 days after these increases in Fe upwelling flux (assuming a 50 m mixed depth and a 40 pM Fe threshold for optimal growth). According to our model, these Fe upwelling fluxes would require only 2–5 nM Fe in the coastal undercurrent feeding the EUC (Table 1, Fig. 2). Extremely intense diatom production, such as occurred during IronEx II ($\sim 300 \text{ mmol C m}^{-2} \text{d}^{-1}$), would require 4–10 nM Fe in the source waters. Although these values are significantly higher than our present estimates for Fe in PNG undercurrents, they are similar to those measured in subsurface waters off the California shelf¹⁶. We now consider whether it is reasonable to expect such increases to have occurred in the past.

Papua New Guinea straddles a complex island arc junction of several plates¹⁷. During the Neogene period (23–1.6 Myr ago) there was significant crustal shortening and uplift in the New Guinea mobile belt, with extensive igneous activity occurring in the New Guinea Highlands during the middle Miocene epoch and again in the Pliocene epoch¹⁸. Between 16 to 3 Myr ago an extensive turbidite unit (called the Sukurum Formation), composed of volcanic and volcanoclastic rocks, was deposited along the northern margin of PNG, derived mainly from the Highlands region (Fig. 3)¹⁹. The Highlands uplift intensified dramatically during the late Miocene

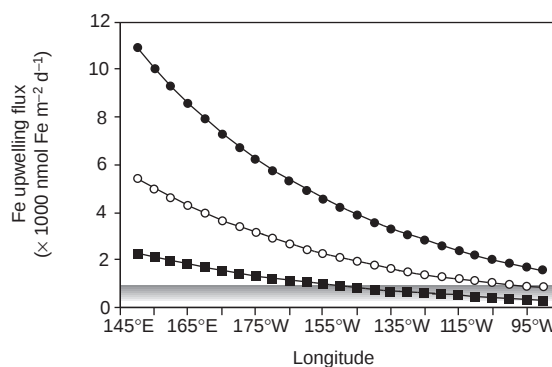


Figure 2 Estimated upwelling fluxes of iron. Iron fluxes from the EUC to surface waters vary as a function of longitude along the Equator and different starting Fe concentrations in the coastal undercurrents of PNG (filled squares, 2.1 nM Fe; open circles, 5.0 nM Fe; filled circles, 10.0 nM Fe). Shaded region indicates the range of upwelling Fe flux needed to support or surpass nitrate-sustained diatom blooms across the eastern equatorial Pacific (Table 1).

and early Pliocene (3–8 Myr ago)²⁰, along with uplift of the north-west margin of PNG resulting from the oblique collision with the New Guinea Highlands²¹. At approximately the time of this collision, the Manus basin began opening, the New Britain arc reversed itself, and the crust under Solomon Sea began subducting beneath the south margin of New Britain²².

Deep waters existed along the northern margin of PNG during this tectonic rearrangement²³, so it is reasonable to expect that the EUC source existed as a single shallow undercurrent constrained against the evolving PNG shelf by coriolis force, much as the New Guinea Coastal Undercurrent is today. Thus the undercurrent would have laid directly over the accumulating Sukurum Formation. The dramatic increase in erosion during the late Miocene to Pliocene (3–8 Myr ago) probably influenced undercurrent Fe concentrations through the partial dissolution of sinking particulates carried offshore in estuarine plumes as well as by turbidity currents induced by slumping of rapidly accumulating shelf sediments.

But by the middle Pliocene (3 Myr ago), uplift and erosion of the Highlands had abruptly decreased²⁴. The oblique collision and uplift of the Adelbert-Finisterre terrane above sea level during this interval also would have redirected the bulk of erosional debris towards the resultant embayment¹⁹, and displaced the undercurrent northwards away from the region of most intense sediment dispersal (Fig. 3).

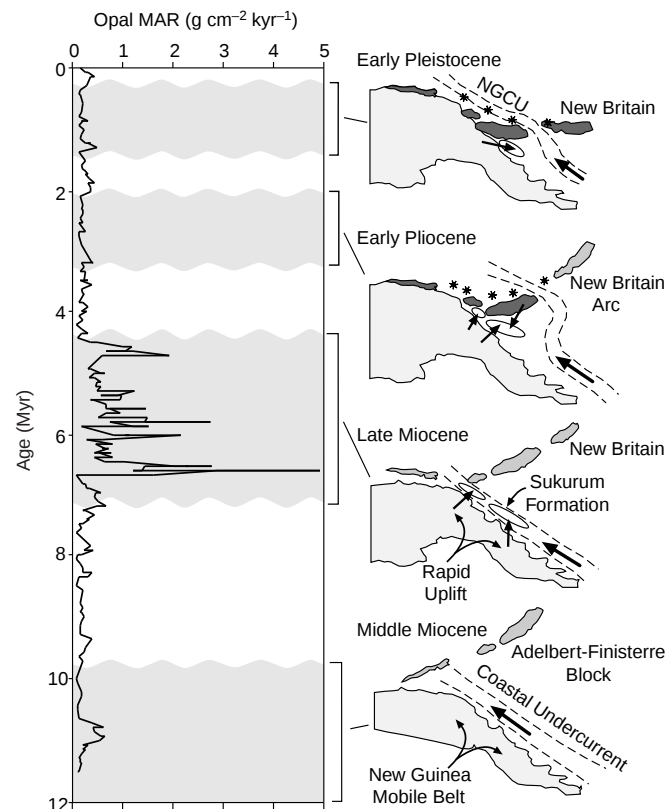


Figure 3 PNG tectonic sequence and eastern equatorial Pacific opal accumulation. Schematic representation of the PNG reconstruction is adapted from Abbott²¹. Right, the submarine (>500 m) Adelbert-Finisterre terrane is shown in light grey while the sub-aerial stages are in dark grey. The palaeo coastal undercurrent feeding the EUC is constrained along the shelf by Coriolis force. Arrows depict the undercurrent direction and the accumulation site of the Sukurum Formation. Asterisks indicate the position of the Bismarck arc comprising numerous volcanic centres. The approximate time line for collision and uplift of the Adelbert-Finisterre terrane is shown. Details of the location and orientation of the submarine terranes are less relevant than the timing of their uplift. Left, the opal mass accumulation rate (MAR) at the eastern equatorial site 849 of the Ocean Drilling Program Leg 138 corresponding to this time interval is shown for comparison³⁰.

During the late Pliocene epoch and the Quaternary period, continued uplift of New Britain, New Ireland and associated island chains of the Manus arc gradually fragmented the palaeo-undercurrent into the present assemblage of undercurrents.

This tectonic reconstruction²¹ suggests that Fe infusion of the EUC has fluctuated over the past 16 Myr, with inputs being most intense during the rapid uplift of the Highlands between ~3 and 8 Myr ago. This time interval corresponds to a period of extremely high opal accumulation in the eastern equatorial Pacific (Fig. 3). Although a paucity of geological data precludes a detailed reconstruction of PNG coastal geometry, the Adelbert-Finisterre terrane uplift was coeval with decreased diatom production (Fig. 3), suggesting a link between tectonics of PNG and high opal accumulation in the eastern equatorial Pacific. The same erosional process would have also infused the undercurrent with other bioactive metals and Si (ref. 25), ensuring that diatoms could flourish in eastern equatorial Pacific surface waters.

Aeolian deposition in the eastern equatorial Pacific would have provided an additional source of Fe, and sediment records indeed show higher deposition occurred between 4 and 7 Myr ago. But these dust fluxes appear to have been only double present inputs²⁶; enough to perhaps stimulate diatom production but not enough to generate massive diatom blooms (Table 1). Moreover, aeolian deposition seems unlikely to explain laminated diatom mats. These features signify very short periods of extremely high diatom production and extend >3,000 km along the Equator beneath the undercurrent⁴ ($\pm 2^\circ$ of the Equator) but not farther north where aeolian fluxes were highest. The essential pulsed Fe inputs which fuelled these very large ($>5 \times 100 \mu\text{m}$) diatoms have left no obvious terrigenous particulate signals.

We now consider the subtle control that erosional inputs may have exerted on the character of algal productivity. Comparatively large Fe enrichments ($>5 \text{ nM}$) of the coastal undercurrent could initiate massive diatom blooms across the entire eastern equatorial region (Table 1), where rapid N depletion would lead to sudden decreases ('crashes') in diatom populations, events that probably caused the formation of laminated diatom mats^{4,27}. These extensive mats appear sporadically between ~4 and 12 Myr ago, which coincides roughly with the Sukurum deposition in PNG. On the other hand, pulses or prolonged periods of moderate Fe enrichment of the coastal undercurrent would stimulate diatom growth over that of their smaller coccolithophore cousins, which could help explain the sub-Milankovitch cycling of silica:carbonate ratios observed in equatorial sediments. Depending on the balance with macronutrient inputs, moderate enrichments also might yield an east to west gradient in diatom production. Conversely, small Fe inputs would stimulate only coccolithophores, because their greater efficiency of Fe use and their smaller size allows them to out-compete diatoms under Fe-stress conditions⁷. Such enrichments might help to explain periods of high carbonate accumulation in equatorial sediments.

Erosion in PNG would probably also have enriched the EUC with the non-bioactive elements Al, Ti, Ba, Th, Po and Pb, which are metals used to help reconstruct palaeoproductivity. It remains to be evaluated whether these chemical proxies may at times have been compromised in the eastern equatorial Pacific.

The potential effect that erosion in the New Guinea Highlands exerted on productivity in the eastern equatorial Pacific would have diminished over the past ~4 Myr, when island emergence north of PNG partitioned the coastal undercurrent feeding the EUC into several smaller undercurrents (Fig. 1). Thus much larger Fe inputs to the New Guinea Coastal Undercurrent should now be required to equal the effects on productivity in the past. Nonetheless, the Bismarck and New Ireland volcanic arcs consist of hundreds of geologically recent eruptive centres and hydrothermal vent regions²⁸, suggesting that tectonics in the PNG region continue to influence sediment records of the eastern equatorial Pacific.

We recognize that climate forcing has unquestionably affected productivity in the eastern equatorial Pacific, and that equatorial sediment records reflect these impacts. But we argue that some aspect of this cycling has been driven by localized tectonic processes in the PNG island-arc region. The challenge is now to determine which patterns are linked to which driving mechanisms. □

Methods

Model description. A first-order approximation of the Fe loss due to upwelling from the EUC during its passage across the Pacific was given by assuming the advective upwelling flux of Fe is ϕw , where ϕ is the dissolved Fe concentration and w the vertical velocity. For a given initial Fe concentration in EUC source waters, the gradient of ϕ is calculated using $\partial\phi/\partial t = -\partial(w\phi)/\partial z$ by supposing that w is uniform across the Pacific ($\sim 1.23 \text{ m d}^{-1}$)²⁹ and Fe concentration falls linearly to zero over a vertical distance δz of 100 m; that is, from the core of the undercurrent at 150 m, across the undercurrent interface, to 50 m in the surface layer⁹. The resulting exponential fall in EUC dissolved Fe concentrations during its passage across the Pacific is described by $\phi = \phi_0 \exp(-x/x_0)$, where x is distance eastwards, $x_0 = u\delta z/w$, and u is an assumed mean undercurrent velocity ($\sim 1 \text{ m s}^{-1}$). Both w and u can be expected to vary somewhat, but the values chosen probably overestimate Fe depletion rates in the western equatorial Pacific, where upwelling velocities should be lower and undercurrent velocities higher.

Received 16 November 1998; accepted 16 February 1999.

- Coale, K. H. *et al.* A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific. *Nature* **383**, 495–501 (1996).
- Martin, J. H. Glacial–interglacial CO₂ change: the iron hypothesis. *Paleoceanography* **5**, 1–13 (1990).
- Pisias, N. G., Mayer, L. A. & Mix, A. C. Paleoceanography of the eastern equatorial Pacific during the Neogene: synthesis of Leg 138 drilling results. *Proc. ODP Sci. Res.* **138**, 5–24 (1995).
- Kemp, A. E. S. & Baldauf, J. G. Vast Neogene laminated diatom mat deposits from the eastern equatorial Pacific Ocean. *Nature* **362**, 141–144 (1993).
- Martin, J. H. *et al.* Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature* **371**, 123–129 (1994).
- Hudson, R. J. M. & Morel, F. M. M. Iron transport in marine phytoplankton: Kinetics of cellular and medium coordination reactions. *Limnol. Oceanogr.* **35**, 1002–1020 (1990).
- Sunda, W. G. & Huntsman, S. A. Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Mar. Chem.* **50**, 189–206 (1995).
- Wells, M. L., Price, N. M. & Bruland, K. W. Iron chemistry in seawater and its relationship to phytoplankton: a progress report. *Mar. Chem.* **48**, 157–182 (1995).
- Coale, K. H., Fitzwater, S. E., Gordon, R. M., Johnson, K. S. & Barber, R. T. Control of community growth and export production by upwelled iron in the equatorial Pacific Ocean. *Nature* **379**, 621–624 (1996).
- Tsuchiya, M., Lukas, R., Fine, R. A., Firing, E. & Lindstrom, E. Source waters of the Pacific equatorial undercurrent. *Prog. Oceanogr.* **23**, 101–147 (1989).
- Butt, J. & Lindstrom, E. Currents off the east coast of New Ireland, Papua New Guinea, and the relevance to regional undercurrents in the western Equatorial Pacific Ocean. *J. Geophys. Res.* **99**, 12503–12514 (1994).
- Philander, S. G. H., Hurlin, W. J. & Seigel, A. D. Simulation of the seasonal cycle of the tropical Pacific Ocean. *J. Phys. Oceanogr.* **17**, 1986–2002 (1987).
- Martin, J. H., Gordon, R. M., Fitzwater, S. & Broenkow, W. W. Vertex: phytoplankton/iron studies in the Gulf of Alaska. *Deep Sea Res.* **36**, 649–680 (1989).
- Murray, J. W., Barber, R. T., Roman, R. R., Bacon, M. P. & Feely, R. A. Physical and biological controls on carbon cycling in the equatorial Pacific. *Science* **266**, 58–65 (1994).
- Takeda, S. & Obata, H. Response of equatorial Pacific phytoplankton to subnanomolar Fe enrichment. *Mar. Chem.* **50**, 219–227 (1995).
- Landing, W. M. & Bruland, K. W. The contrasting biogeochemistry of iron and manganese in the Pacific Ocean. *Geochim. Cosmochim. Acta* **51**, 29–43 (1987).
- Tregoning, P. *et al.* Estimation of current plate motions in Papua New Guinea from Global Positioning System observations. *J. Geophys. Res.* **103**, 12181–12203 (1998).
- Rogerson, R. J. *et al.* The Geology and Mineral Resources of the Sepik Headwaters Region, Papua New Guinea (Geological Survey of Papua New Guinea, Port Moresby, 1987).
- Abbott, L. D. *et al.* Stratigraphic constraints on the development and timing of arc–continent collision in northern Papua–New Guinea. *J. Sedim. Res. B* **64**, 169–183 (1994).
- Hill, K. C. & Gleadow, A. J. W. Uplift and thermal history of the Papuan fold belt, Papua New Guinea: apatite fission track analysis. *Austr. J. Earth Sci.* **36**, 515–539 (1989).
- Abbott, L. D. Neogene tectonic reconstruction of the Adelbert–Finisterre New Britain collision, northern Papua New Guinea. *J. S. Asian Earth Sci.* **11**, 33–51 (1995).
- Musgrave, R. J. Paleomagnetism and tectonics of Malaita, Solomon Islands. *Tectonics* **9**, 735–759 (1990).
- Abbott, L. D. *et al.* Measurement of tectonic surface uplift rate in a young collisional mountain belt. *Nature* **385**, 501–507 (1997).
- Crowhurst, P. V., Hill, K. C. & Foster, D. A. in *Proc. PNG Geology, Exploration and Mining Conf.* (ed. Hancock, G. E.) 51–60 (Australian Inst. of Mining and Metallurgy, Melbourne, 1997).
- Dugdale, R. C. & Wilkerson, F. P. Silicate regulation of new production in the equatorial Pacific upwelling. *Nature* **391**, 270–273 (1998).
- Hovan, S. A. Late Cenozoic atmospheric circulation intensity and climatic history recorded by eolian deposition in the eastern equatorial Pacific Ocean, Leg 138. *Proc. ODP Sci. Res.* **138**, 615–625 (1995).
- Kemp, A. E. S., Baldauf, J. G. & Pearce, R. B. Origins and paleoceanographic significance of laminated diatom ooze from the eastern equatorial Pacific Ocean. *Proc. ODP Sci. Res.* **138**, 641–646 (1995).
- Auzende, J.-M. *et al.* Activité tectonique, magmatique et hydrothermale dans le bassin de Manus (SW Pacifique, Papouasie–Nouvelle Guinée): campagne MANUSFLUX du Shinkai-6500. *CR Acad. Sci. Paris* **323**, 501–508 (1996).

- Bryden, H. L. & Brady, E. C. Diagnostic model of the three-dimensional circulation in the upper equatorial Pacific Ocean. *J. Phys. Oceanogr.* **15**, 1255–1273 (1985).
- Farrell, J. W. *et al.* Late Neogene sedimentation patterns in the eastern equatorial Pacific Ocean. *Proc. ODP Sci. Res.* **138**, 717–753 (1995).

Acknowledgements. We thank L. D. Abbott, T. F. Pedersen and J. R. Toggweiler for comments. This work was supported by the NSF Chemical Oceanography Program.

Correspondence and requests for materials should be addressed to M.L.W. (e-mail: mlwells@maine.edu).

Partitioning of nickel and cobalt between silicate perovskite and metal at pressures up to 80 GPa

O. Tschauner*, A. Zerr*, S. Specht*, A. Rocholl†, R. Boehler* & H. Palme‡

* Max-Planck-Institut für Chemie, Postfach 3020, 55020 Mainz, Germany

† Mineralogisches Institut, Universität Heidelberg, Im Neuenheimer Feld 236, 69120 Heidelberg, Germany

‡ Institut für Mineralogie und Geochemie, Universität Köln, Zùlpicher Straße 49b, 50674 Köln, Germany

The high abundance of both nickel and cobalt and the chondritic Ni/Co ratio found in samples derived from the Earth's mantle are at odds with results from laboratory-based partitioning experiments conducted at pressures up to 27 GPa (refs 1,2). The laboratory results predict that the mantle should have a much lower abundance of both Ni and Co and a considerably lower Ni/Co ratio owing to the preferential partitioning of these elements into the iron core. Two models have been put forward to explain these discrepancies: homogeneous accretion^{3–6} (involving changes of the Ni and Co partition coefficients with oxygen and sulphur fugacities, pressure and temperature) and heterogeneous accretion^{7–9} (the addition of chondritic meteorites to the mantle after core formation was almost complete). Here we report diamond-cell experiments on the partitioning of Ni and Co between the main lower-mantle mineral ((Mg,Fe)SiO₃-perovskite) and an iron-rich metal alloy at pressures up to 80 GPa (corresponding to a depth of $\sim 1,900$ km). Our results show that both elements become much less siderophilic with increasing pressure, such that the abundance of both Ni and Co and the Ni/Co ratio observed in samples derived from the Earth's mantle appear to indeed be consistent with a homogeneous accretion model.

Heterogeneous accretion models have garnered much support as an explanation of the high abundance of Ni and Co in the Earth's upper mantle¹⁰. This support is based on studies on the partitioning of these elements between metal and silicate phases up to pressures of 27 GPa (refs 11–16) which did not show convergence to equal partitioning of Ni and Co as required in order to preserve their observed chondritic ratio. However, studies^{5,6} on liquid metal–silicate systems showed that this chondritic ratio may be preserved under certain conditions of pressure and temperature in an early, sulphur-rich magma ocean. None of these studies, however, addressed the partitioning behaviour of Ni and Co at lower-mantle conditions experimentally. About 80% of the lower mantle consists of (Mg,Fe)SiO₃-perovskite (hereafter called perovskite). Because of the high melting temperature of this phase at high pressure¹⁷, a substantial amount of solid perovskite was probably present in the early Earth during core formation. Models for the interaction between iron and perovskite are nearly unconstrained by experimental data: only few data exist on perovskite/metal partitioning of Ni and Co at pressures of the uppermost lower mantle^{11,12}. The chemical exchange of siderophile elements between perovskite and metal at very high pressure is also important for understanding

possible core–mantle interactions. We therefore performed the first, to our knowledge, diamond-cell experiments on the partitioning of Ni and Co between an FeNiCo-alloy and perovskite at pressures appropriate to most of the Earth's lower mantle.

The two biggest challenges in performing these experiments were to ensure chemical equilibrium at the extreme pressure and temperature conditions in the laser-heated diamond cell and to analyse accurately the small samples with sub-micrometre spatial resolution. To provide the minimal pressure and temperature gradients necessary to achieve equilibrium we used inert and nearly hydrostatic pressure media of low thermal conductance, thus providing very low temperature gradients across the samples heated with high-power CO₂ lasers. To achieve submicrometre analytic resolution we used the depth profiling capabilities of the ion probe secondary ion mass spectrometry (SIMS) analysis.

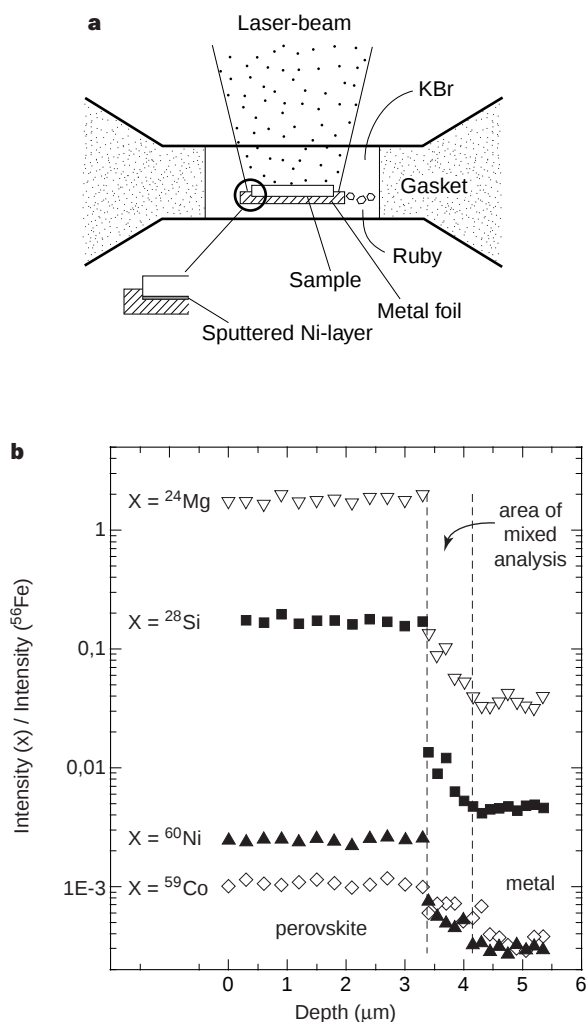


Figure 1 Experimental and analytical details. **a**, Sketch of the sample assemblage in the diamond cell (see text for details). **b**, Analytical depth profile from SIMS analysis of a sample assemblage (D56), starting at the perovskite surface showing measured intensities of ²⁴Mg, ²⁸Si, ⁵⁹Co and ⁶⁰Ni relative to ⁵⁶Fe. These intensities depend on element sensitivity concentration and matrix. For example, relatively light masses like ²⁴Mg and ²⁸Si yield higher count rates than a relatively heavier one like ⁶⁰Ni, and count rates are significantly smaller in a metal matrix than in a silicate matrix. The Mg and the Si content in the metal in the present sample are ~10 p.p.m. and 1.5 at.%, respectively, while the Ni and the Co contents are 0.7 and 0.2 at.% in the metal, and ~0.2 at.% in the perovskite, respectively. The area of mixed analysis is ~1 μm, if the perovskite and the metal parts of a sample could be separated (see text); no mixed analysis occurred. The concentrations (Table 1) are calculated from the shown intensities by conversion with standards.

A detailed description of the experimental set-up (Fig. 1a) and the analytical procedure is given in Methods. We chose a temperature of $2,100 \pm 100$ K for the experiments, which is probably near the average temperature of the lower mantle. The pressure range was from 23 to 78 GPa, which corresponds to depths of 650 to 1,850 km in the Earth. The run conditions and the compositions of the run products are listed in Table 1a. The partition coefficients D^X calculated from these compositions are listed in Table 1b ($D^X = X(\text{metal})/X(\text{perovskite})$, where X is the concentration of Fe, Ni and Co in wt%).

Figure 2 shows the pressure dependence of these partition coefficients. D^{Ni} and D^{Co} decrease from values of ~20 at about 26 GPa to values below 5 at pressures higher than 50 GPa. D^{Fe} ranges from 20 to 80 without showing a systematic pressure dependence. The error bars shown in Fig. 2 represent the uncertainties in the

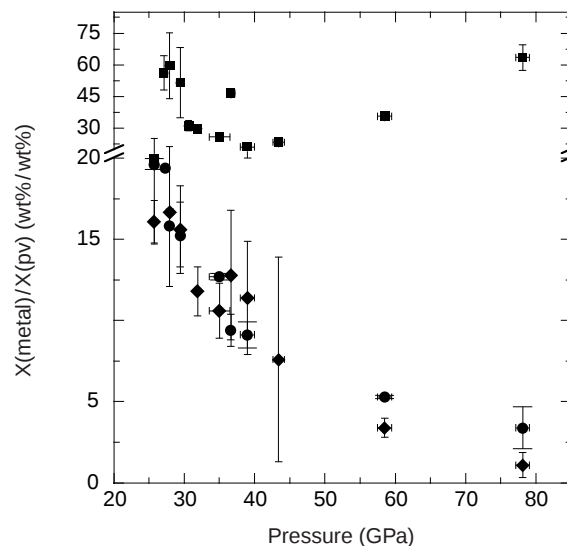


Figure 2 Partition coefficients of Ni (circles), Co (diamonds) and Fe (squares) between metal and perovskite as a function of pressure at $2,100 \pm 100$ K. Partition coefficients from runs with perovskite containing >3 at.% Ni were omitted as marked changes in the Raman spectra of such samples indicated substantial deviation from ideal solution.

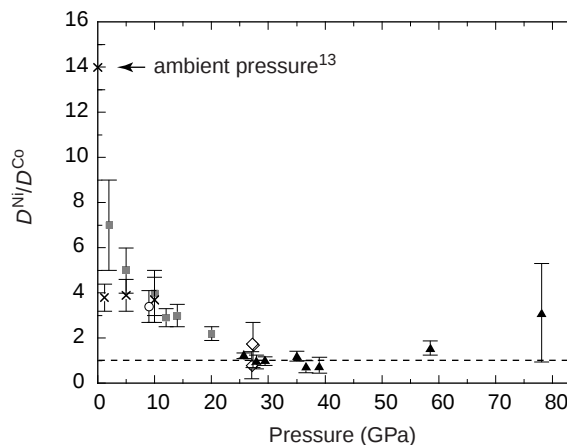


Figure 3 Ratios $D^{\text{Ni}}/D^{\text{Co}}$ from present work and from previous studies for metal-oxide systems relevant to core formation. This ratio is the most suitable for comparing different studies on the partitioning of Ni and Co, because it is nearly independent of temperature within the relevant temperature range (2000–2300 K)^{13,18,20}, and it is nearly independent of f_{O_2} (refs 13, 20). Triangles, present data; diamonds, perovskite according to Ohtani *et al.*^{11,12}. Circle, magnesio-wüstite¹⁸; squares and crosses, liquid silicate^{5,13,19,24}. The dotted line represents the ratio required for models of homogeneous accretion.

chemical analysis. These uncertainties are discussed in Methods. Other possible sources of uncertainty include disequilibrium and variations in the oxygen fugacities. If equilibrium had not been achieved in the experiments, the values of D^X would be even lower than those observed, thus increasing the effect of pressure on the partitioning behaviour of Ni and Co. However, there are several indications that equilibrium was at least approached: (1) the flatness of the concentration profiles obtained through SIMS depth profiling (see Fig. 1b), (2) a successful reversal experiment for Ni, and (3) estimates of diffusion constants for Ni and Co which show that the run durations were long enough. These three issues are discussed in detail in Methods. In addition to disequilibrium, differences in the oxygen fugacities, f_{O_2} , can affect D^X . From the variation in D^{Fe} we conclude that f_{O_2} indeed varies from run to run, but that this variation affects the values of D^{Ni} and D^{Co} less strongly than

pressure. For example, the values of D^{Fe} at the highest pressures are similar to those at the lowest pressures, but the D -values of both Ni and Co decrease by more than a factor of 3. This shows that the decrease in the partition coefficients of Ni and Co is due to pressure, and not a result of changing f_{O_2} . From these observations we conclude that the siderophile character of Ni and Co decreases strongly with pressure. Studies on Ni, Co-partitioning between metal, silicate liquids and magnesiowüstite at lower pressures showed that the ratio D^{Ni}/D^{Co} is nearly independent of oxygen fugacity and temperature^{13,18,19}. Figure 3 shows this ratio as a function of pressure for our data, along with previous results on Ni- and Co-partitioning. This plot, which includes results on perovskite, magnesiowüstite and silicate liquids shows good agreement between previous^{11,12} and present results on perovskite.

One of the principal motivations behind heterogeneous accretion

Table 1 Experimental details and results

(a) Compositions and run conditions

Run conditions				Composition (wt%)									
Expt. no.	P (GPa)	T (K)	t (s)	Metal				Perovskite					
				Fe	Ni	Co	Si*	Mg	Fe	Ni	Co	Si	
D2†	36.8 ± 0.8	>2,000	300	NA††	NA	NA	NA	19.90 ± 0.9 20.42 ± 2.0	1.70 ± 0.03 2.0 ± 1.0	0.58 ± 0.06 0.5 ± 0.2	NA	19.42 ± 0.08 19.07 ± 2.0	
D32‡§	25.7 ± 0.2	2,050 ± 50	300	57.71 ± 8.40	25.48 ± 0.43	7.00 ± 0.41	~10	16.62 ± 0.63	2.95 ± 0.30	1.30 ± 0.12	0.43 ± 0.01	18.70 ± 0.47	
D46§	27.1 ± 0.3	2,000 ± 30	300	67.40 ± 2.17	4.16 ± 0.76	13.18 ± 1.41	~10	18.32 ± 0.70	1.20 ± 0.12	0.30 ± 0.01	0.48 ± 0.01	19.68 ± 0.60	
D17	27.3 ± 0.2	2,100 ± 100	300	97 ± 1	3.5 ± 0.5	—	—	18.76 ± 0.40	1.24 ± 0.03	0.18 ± 0.01	NA	19.80 ± 0.60	
D47§	27.9 ± 0.5	2,040 ± 30	900	85.18 ± 0.73	6.24 ± 0.38	8.58 ± 0.35	0.07 < 0.01	18.06 ± 5.75	1.43 ± 0.24	0.39 ± 0.05	0.51 ± 0.12	19.59 ± 5.35	
D54	29.4 ± 0.5	2,050 ± 50	1,800	87.71 ± 1.08	2.20 ± 0.14	10.09 ± 0.94	>5	17.65 ± 4.50	1.70 ± 0.40	0.15 ± 0.01	0.65 ± 0.05	19.83 ± 4.91	
D25	30.6 ± 0.5	2,070 ± 50	300	65.42 ± 0.28	3.50 ± 0.06	21.98 ± 0.22	>1	14.75 ± 1.90	2.10 ± 0.15	8 ± 3	3.15 ± 0.78	16.37 ± 0.75	
D23	31.9 ± 1.0	1,990 ± 50	1,200	76.27 ± 0.24	21.0 ± 0.2	2.73 ± 0.22	0.19 ± 0.02	17.21 ± 0.10	2.56 ± 0.04	3.3 ± 0.8	0.23 ± 0.01	18.86 ± 0.51	
D41†§	35.0 ± 1.5	2,220 ± 100	150	70.58 ± 1.73	20.60 ± 1.11	8.82 ± 0.61	NA	16.45 ± 0.50	2.72 ± 0.11	1.62 ± 0.18	0.83 ± 0.06	18.36 ± 2.0	
D34†	36.6 ± 0.6	2,240 ± 40	180	90.81 ± 1.06 88.7 ± 1.0	9.01 ± 0.76 6.53 ± 0.41	0.18 ± 0.03 0.27 ± 0.06	0.0043 ± 3 × 10 ⁻⁴	18.04 ± 0.03	1.95 ± 0.02	0.96 ± 0.02	0.014 ± 0.004	19.02 ± 0.02	
D39	39.0 ± 1.0	2,170 ± 50	220	55.91 ± 3.13	36.50 ± 0.93	7.59 ± 2.20	NA	16.57 ± 0.04	2.77 ± 0.01	4.00 ± 0.25	0.66 ± < 0.01	15.98 ± 0.03	
D37	43.4 ± 0.8	2,220 ± 50	400	84.0 ± 0.5	11.3 ± 0.45	4.72 ± 0.39	0.031 ± 0.011	17.27 ± 1.4	2.11 ± 0.01	5.99 ± 1.07	0.62 ± < 0.01	14.00 ± 1.20	
D28	58.5 ± 1.0	1,970 ± 30	300	87.7 ± 2.07	10.5 ± 1.91	2.16 ± 0.16	0.17 ± 0.01	16.91 ± 0.60	2.46 ± 0.03	1.97 ± 0.01	0.63 ± 0.07	18.01 ± 0.55	
D56	78.1 ± 1.0	2,000 ± 40	480	99.12 ± 0.27	0.70 ± 0.20	0.18 ± 0.07	1.5 ± 0.1	18.27 ± 1.24	1.56 ± 0.13	0.20 ± 0.02	0.17 ± 0.05	19.78 ± 1.17	
E12#	30.5 ± 0.5	2,250 ± 70	180	—	—	—	—	15.09 ± 0.59	1.86 ± 0.12	0.042 ± 0.003	3.05 ± 0.007	19.93 ± 0.39	
—II	~120	3,000–3,500	1800	99.97 ± 0.01	NA	NA	0.030 ± 0.003	15.3 ± 2.0	4.7 ± 0.7	NA	NA	19.98 ± 2.0	

(b) Partition coefficients

Run conditions				Partition coefficients		
Nr	P (GPa)	T (K)	t (s)	D^{Ni}	D^{Co}	D^{Fe}
D32	25.7 ± 0.2	2,050 ± 50	300	19.63 ± 0.34	16.1 ± 1.3	19.63 ± 0.53
D46	27.1 ± 0.3	2,000 ± 30	300	13.93 ± 0.30	27.5 ± 3.5	56.2 ± 0.82
D17	27.3 ± 0.2	2,200 ± 100	300	19.4 ± 0.2	NA	78 ± 2
D47	27.9 ± 0.5	2,040 ± 30	900	15.83 ± 0.30	16.7 ± 4.6	59.6 ± 1.56
D54	29.4 ± 0.5	2,050 ± 50	1,800	15.3 ± 2.0	15.6 ± 2.7	51.6 ± 1.67
D25	30.6 ± 0.5	2,070 ± 50	300	0.4 ± 0.3	15.9 ± 1.8	31.2 ± 0.25
D23	31.9 ± 1.0	1,990 ± 50	1,200	6.4 ± 1.5	11.8 ± 1.5	29.8 ± 0.4
D41	35.0 ± 1.5	2,220 ± 100	150	12.7 ± 2.1	10.6 ± 1.7	25.9 ± 1.8
D34	36.6 ± 0.6	2,240 ± 40	180	9.4 ± 1.0	12.8 ± 4.0	46.6 ± 1.0
D39	39.0 ± 1.0	2,170 ± 50	220	9.12 ± 0.80	11.4 ± 3.5	20.2 ± 1.2
D37	43.4 ± 0.8	2,220 ± 50	400	7.3 ± 0.5	7.60 ± 0.63	39.8 ± 0.4
D28	58.5 ± 1.0	1,970 ± 30	300	5.3 ± 1.0	3.4 ± 0.6	35.7 ± 1.1
D56	78.1 ± 1.0	2,000 ± 40	480	3.4 ± 1.3	1.09 ± 0.77	63.5 ± 6.0

We used single crystals of orthopyroxene from San Carlos and from an antarctic periodotite with compositions (Mg_{0.817}, Fe_{0.089}, Ca_{0.018}) (Si_{0.924})O₃ with 0.108% Al and (Mg_{0.81}, Fe_{0.102}, Ca_{0.017}, Si_{0.945})O₃ with 0.006% Al. The metallic starting materials were FeNiCo-alloy foils of average composition Fe_{0.56}Ni_{0.28}Co_{0.176}. The amounts of metal and silicate ranging between 10 and 100 ng were estimated from their volumes. The amount of metal was at least three times of that of the silicate. Experiments D34, 46, 47, 56 were performed using a NaCl pressure medium, in D2 and D17 we used Ar as pressure medium. The analysis was performed with a primary ion current of 0.7 to 0.9 nA and 12.5 kV accelerating voltage. The imaged area was 25 μm, the beam spot was 4 μm FWHM and rastered over 10 × 10 μm². The accelerating voltage of the secondary ions was 4.5 kV. We set an energy window of 17 eV. No energy offset was used. The mass resolution was 4100 to 4500 M/dm.

* The Si content of the metal was evaluated by using Si-containing meteoritic iron. In some cases the metal phase in the run products of the present experiments contained regions of ~100 nm diameter which were enriched in K, Ca, Si, Fe, Mn and O as evident from changes in the sensitivities of Fe and Si. These regions may result from exsolution during quench. Similar phenomena have been reported from multi-anvil experiments on metal/silicate systems (for example, ref. 18).

† We cross-checked the SIMS analysis of the perovskite by SEM-EDX on separate, thicker samples heated in contact with a small amount of Ni, sputtered on one surface. This resulted in droplets on the silicate surface containing both Fe and Ni, and thus allowing separate analysis of the silicate taking account of secondary excitations of Fe and Ni in the metallic droplets. We give the analyses of D2 as an example. The EDX analysis is in italics. Microprobe analysis of samples as thin as 3 μm usually fails, but we were able to analyse the metal sample of run D34 with microprobe and SIMS successfully. The microprobe analysis is in italics.

‡ We used the perovskite sample of run E12 as starting material in run 32. In run 41 we used the perovskite sample of run D28 as starting material.

§ Perovskite as starting material.

|| In this experiment a foil of pure iron was embedded in MgSiO₃ gel which filled the rest of the cell chamber (see Fig. 1). The metal was heated close to the melting point. Afterwards the sample was polished. Adjacent to the metal the gel was transformed to perovskite.

†† Not analysed.

No metal phase present.

models was the partitioning behaviour of Ni and Co at low pressure^{1,2,8,18}, which is in contrast to the observed high abundance and nearly chondritic ratio of these elements in upper-mantle rocks. But these arguments are not valid if chemical exchange between metal and silicate at lower-mantle pressures was important during core formation, because metal-perovskite partitioning does not fractionate Ni from Co and would provide a high abundance of both elements in the mantle. Although the metal was probably molten during core formation, our results for solid metal are applicable, because the liquid-solid iron partitioning of Ni and Co is close to unity^{20–22}. If the bulk Earth is chondritic with respect to refractory elements and if the mantle (with a Ni content²³ of 0.2 wt%) is chemically homogeneous, the present partition coefficients imply drastic chemical disequilibrium between core and mantle. This could result in a chemical boundary layer at the core–mantle interface enriched in Ni by up to 5 wt%, on the basis of recent estimates²³ of the composition of the outer core. □

Methods

Experimental and analytical procedures. The silicate portions of our samples were either single crystals of San Carlos orthopyroxene, or single crystals of perovskite synthesized from orthopyroxene in a CO₂-laser-heated diamond cell, with diameters of ~50 µm, and thicknesses of 5, 10 or 15 µm. These silicate crystals were in contact with FeNiCo-alloy foils. The variation of the Ni and Co contents in the foils was ~±5% on the scale of a few µm. The thickness of the metal foils was 5 µm and their diameters were larger than those of the silicate crystals. A Ni layer of ~20 nm thickness was sputtered on one side of these crystals to improve the contact between silicate and metal. This sandwich assemblage was embedded in a KBr or a NaCl pressure medium. KBr is essentially chemically inert (see Table 1), provides nearly hydrostatic pressure conditions, and has a low thermal conductivity. NaCl was used for several runs, which due to its lower background luminescence is more suitable for *in situ* Raman spectroscopic identification of the perovskite phase. A power-stabilized, defocused CO₂-laser beam²⁴ was used to heat the sample (Fig. 1a). The main absorption of laser radiation occurs in the silicate crystals. The pressure medium does not absorb laser radiation and does not emit incandescent light. The small thickness of the crystals provided negligible axial temperature gradients. The radial temperature variation was measured with a spatial resolution of 0.3 µm, and was found to be within the uncertainty of the temperature measurements (about ±60 K). A more detailed description of the temperature measurement is in ref. 24. Temperature quenching was achieved within milliseconds by switching off the laser. The sample assemblages were recovered by dissolving the pressure medium with twice-distilled water. Owing to some plastic deformation during the run the recovered perovskite samples had thicknesses of ~5 µm, and the metal samples were 2–3 µm thick.

Because electron microprobe analysis is unsuitable owing to secondary X-ray emission, we used the ion micro-probe SIMS technique, which provides high analytical sensitivity and a sub-micrometre depth resolution, for the chemical analysis of the perovskite and the adjacent metal. We used the Cameca 3f ims ion probes at the MPI, Mainz and at the Mineralogical Institute, Heidelberg. Depth profiling was started either at the metal–silicate contact by mechanically separating both phases or on the entire sample assemblage (Fig. 1b). Each depth profile consisted of at least 30 analyses. Two depth profiles were taken for each sample. The analytical sensitivity of the SIMS measurements in this study was 1×10^{-3} at.% Ni and 1×10^{-4} at.% Co in the silicate, and 0.1 at.% in the metal. Quantification of the chemical composition requires the conversion of measured mass ratios to concentrations of elements using standards which have a similar matrix. This means that the standards have to be closely related to the samples both in structure and composition. This conversion is the main source of systematic errors. We produced standards of various compositions by converting orthopyroxene to perovskite in the CO₂-laser-heated diamond cell. The starting material was analysed by both micro-probe and neutron activation. SIMS analysis of separate perovskite samples using these standards was then crosschecked with a microprobe (Table 1a). The composition of the metal in the partition experiments was determined in a similar way, using alloys of known compositions as standards and by cross-

checking SIMS analysis with a microprobe (Table 1). The crosschecks show that systematic errors were within the error bars.

The concentration profiles of the samples are flat in all experiments except in two pressure regimes, from 23 to 25 and from 29 to 32 GPa, where we found gradients in the Co-, Ni- and Fe-concentrations of the perovskite samples—even at run durations higher than 1,000 s. We believe that this strong decrease in diffusion in these small pressure ranges is due to as-yet unknown transitions between different perovskite phases. We omitted all experiments from these two pressure regions where the gradients were larger than 10% of the mean values. However, we used these concentration gradients to estimate the diffusion constants of Ni and Co, resulting in values between 5×10^{-12} and 1×10^{-13} m² s⁻¹. Using the approximation $\Delta x = \sqrt{Dt}$, with Δx = distance, D = diffusion constant and t = time, and assuming $t = 300$ s as a typical run duration, we find values of Δx ranging from 15 to 30 µm. The thickness of our perovskite samples was ~5 µm, and thus equilibrium was probably achieved.

In all our experiments Ni and Co diffused out of the metal into the silicate; but Fe diffused in the opposite direction. In addition, we performed two reversal experiments of Ni and Co (D41 and D32) using perovskite samples as starting material which were enriched either in Ni or Co. Run D41 was successful for Ni: a comparison to a normal (unreversed) run at similar pressure and temperature, and with similar D^{Fe} s yielded comparable D^{Ni} s within the error bars (runs D41 and D34 in Table 1a). For the case of Co, however, comparison of a reversed run (D32) to unreversed runs is ambiguous because of strong differences in D^{Fe} .

Received 30 March 1998; accepted 23 February 1999.

- Ringwood, A. E. *Origin of the Earth and Moon* 39 (Springer, New York, 1979).
- Schmitt, W., Palme, H. & Wänke, H. Experimental determination of metal/silicate partition coefficients for P, Co, Ni, Cu, Ga, Ge, Mo, W. *Geochim. Cosmochim. Acta* **53**, 173–185 (1989).
- Ringwood, A. E. *Origin of the Earth and Moon* 122–134 (Springer, New York, 1979).
- Murthy, R. V. Early differentiation of the Earth and the problem of mantle siderophile elements: A new approach. *Science* **253**, 303–306 (1991).
- Li, J. & Agee, C. B. Geochemistry of mantle-core formation at high pressure. *Nature* **381**, 686–689 (1996).
- Righter, K., Drake, M. J. & Yaxley, G. Prediction of siderophile element metal/silicate partition coefficients to 20 GPa and 2800 °C. *Phys. Earth Planet. Inter.* **100**, 115–134 (1997).
- Ringwood, A. E. Chemical evolution of the terrestrial planets. *Geochim. Cosmochim. Acta* **30**, 41–104 (1966).
- Turekian, K. & Clark, S. P. Inhomogeneous accumulation of the Earth from the primitive solar nebula. *Earth Planet. Sci. Lett.* **6**, 346–348 (1969).
- Wänke, H. Chemical composition and accretional history of terrestrial planets. *Phil. Trans. R. Soc. Lond. A* **325**, 545–557 (1988).
- O'Neill, H. St. C. & Palme, H. In *The Earth's Mantle* (ed. Jackson, I.) 3–123 (Cambridge Univ. Press, 1998).
- Ohtani, E., Yurimoto, H. & Seto, S. Element partitioning between metallic liquid, silicate liquid, and lower mantle minerals: Implications for core formation of the Earth. *Phys. Earth Planet. Inter.* **100**, 97–114 (1997).
- Ohtani, E. & Kato, T. Transition metal partitioning between lower mantle and core materials at 27 GPa. *Geophys. Res. Lett.* **18**, 85–88 (1991).
- Gessmann, C. K. & Rubie, D. C. The effect of temperature on the partitioning of Ni, Co, Mn, Cr and V at 9 GPa and constraints on formation of the Earth's core. *Geochim. Cosmochim. Acta* **62**, 667–682 (1998).
- Thibault, Y. & Walter, M. J. The influence of pressure and temperature on the metal–silicate partition coefficients of nickel and cobalt. *Geochim. Cosmochim. Acta* **59**, 991–1002 (1995).
- Hillgren, V. J., Drake, M. J. & Rubie, D. C. High pressure and high temperature metal–silicate partitioning of siderophile elements: The importance of silicate liquid composition. *Geochim. Cosmochim. Acta* **60**, 2257–2263 (1996).
- Dingwell, D. B., O'Neill, H. St. C., Ertel, W. & Spettel, B. The solubility and oxidation state of Ni in silicate melt at low oxygen fugacity. *Geochim. Cosmochim. Acta* **58**, 1967–1974 (1994).
- Zerr, A. & Boehler, R. Melting of (Mg, Fe)SiO₃-perovskite to 625 kilobars: Indication of a high melting temperature in the lower mantle. *Science* **262**, 553–555 (1993).
- Holzheid, A., Borisov, A. & Palme, H. The effect of oxygen fugacity and temperature on solubilities of nickel, cobalt and molybdenum in silicate melts. *Geochim. Cosmochim. Acta* **58**, 1975–1981 (1994).
- Capobianco, C. J. & Amelin, A. A. Metal–silicate partitioning of nickel and cobalt: the influence of temperature and oxygen fugacity. *Geochim. Cosmochim. Acta* **58**, 125–140 (1994).
- Jones, J. H. & Drake, M. J. Geochemical constraints on core formation in the Earth. *Geochim. Cosmochim. Acta* **47**, 1199–1209 (1986).
- Jones, J. & Malvin, D. J. A nonmetal interaction model for the segregation of trace metals during solidification of Fe–Ni–S, Fe–Ni–P, and Fe–Ni–S–P alloys. *Mettall. Trans. B* **21**, 697–706 (1990).
- Jones, J. & Walker, D. Partitioning of siderophile elements in the Fe–Ni–S system: 1 bar to 80 kbar. *Earth Planet. Sci. Lett.* **105**, 127–133 (1991).
- McDonough, W. F. & Sun, S.-s. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).
- Boehler, R. & Chopelas, A. A new approach to laser heating in high pressure mineral physics. *Geophys. Res. Lett.* **18**, 1147–1150 (1991).

Acknowledgements. We acknowledge discussions with V. J. Hillgren. We thank Th. Ludwig for technical support at the SIMS, A. Diegeler for performing the depth measurements on the samples with AFM, and B. Spettel for INA-analyses of some single crystal standards. This work was partly supported by the priority programme "Elementverteilungen" of the Deutsche Forschungsgemeinschaft (DFG).

Correspondence and requests for materials should be addressed to O.T. (e-mail: Tschau@mpch-mainz.mpg.de).

Simulating the effects of climate change on tropical montane cloud forests

Christopher J. Still^{*‡}, Prudence N. Foster[†]
& Stephen H. Schneider^{*}

^{*} Department of Biological Sciences, Stanford University, Stanford, California 94305-5020, USA

[†] Center for Climate Studies Research, University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153-8904, Japan

[‡] Carnegie Institution of Washington, Department of Plant Biology, 260 Panama Street, Stanford, California 94305-4150, USA

Tropical montane cloud forests are unique among terrestrial ecosystems in that they are strongly linked to regular cycles of cloud formation. We have explored changes in atmospheric parameters from global climate model simulations of the Last Glacial Maximum and for doubled atmospheric carbon dioxide concentration ($2 \times \text{CO}_2$) conditions which are associated with the height of this cloud formation, and hence the occurrence of intact cloud forests. These parameters include vertical profiles of absolute and relative humidity surfaces, as well as the warmth index¹, an empirical proxy of forest type. For the glacial simulations, the warmth index and absolute humidity suggest a downslope shift of cloud forests that agrees with the available palaeodata. For the $2 \times \text{CO}_2$ scenario, the relative humidity surface is shifted upwards by hundreds of metres during the winter dry season when these forests typically rely most on the moisture from cloud contact. At the same time, an increase in the warmth index implies increased evapo-transpiration. This combination of reduced cloud contact and increased evapo-transpiration could have serious conservation implications, given that these ecosystems typically harbour a high proportion of endemic species and are often situated on mountain tops or ridge lines.

Tropical montane cloud forests (TMCs) occur where mountains are frequently enveloped by tradewind-derived orographic clouds and mist in combination with convective rainfall. Many features of these forests are directly or indirectly related to cloud formation, from vegetation morphology to nutrient budgets to solar insolation². One of the most direct impacts of frequent cloud cover is the deposition of cloud droplets through contact with soil and vegetation surfaces (horizontal precipitation)³. Total horizontal precipitation is greater than that from vertical rainfall events in some systems during the dry season, when these forests can experience water stress^{2,4}. Because the combination of horizontal precipitation and lowered evapo-transpiration due to frequent cloud contact significantly increases precipitation minus evaporation in these forests, they function as important local and regional watersheds. Also, owing to the sponge-like effect of epiphytes (for example, mosses, bromeliads and ferns), these forests act as capacitors in regulating the seasonal release of precipitation, thereby providing flood and erosion control in the rainy season and water storage in the dry season.

In addition to their hydrological importance, these ecosystems typically harbour an impressive array of plants and animals. Although the biodiversity of TMCs is not as high as that of lowland moist tropical forests⁵, the level of endemism found in resident animal species is exceptional³. For example, 32% of Peruvian endemic vertebrates are localized in cloud forests⁶. The conservation status of these unique ecosystems is precarious as they are among the most endangered of all tropical forest types. A high annual deforestation rate in tropical mountain forests caused by harvesting fuel wood, resource logging and agricultural conversion is increasingly threatening cloud forests worldwide⁷.

Palaeoclimatic pollen evidence strongly suggests a downslope shift of the range of some current cloud forest species during the last glacial period. There is abundant evidence⁷ for downslope migrations of South and Central American montane taxa (*Quercus*, *Alnus*, *Weinmannia* and *Podocarpus*, for example) during glacial times. *Weinmannia* is now a characteristic genus of cloud forest trees in the high Andes, and *Podocarpus* is found in cloud forests throughout the tropics. Other evidence, ranging from noble gas concentrations in groundwater⁸ to Barbados corals⁹, to snowline depression¹⁰, indicates that certain regions of the tropics were cooler by some 2–5 °C, with considerable variation in both temperature and moisture conditions^{11,12}. Such changes surely affected both the altitudinal and latitudinal distributions of cloud forests in the glacial past.

Cloud formation associated with trade winds often occurs as a result of orographic effects. For example, in Costa Rica the humid trade winds from the Caribbean Sea quickly encounter the continental divide formed by the Cordillera de Tilarán in the north-central portion of the country and the prevailing winds undergo orographic uplift¹³. As these air parcels are uplifted, they expand and cool until their water vapour pressure exceeds their saturation vapour pressure and condensation can occur. The clouds formed in this process are referred to as orographic clouds, and the height at which water vapour condenses is the lifting condensation level (LCL).

This altitude is a function of the lapse rate, which decreases from a dry adiabatic rate to a moist adiabatic rate when condensation occurs during rapid ascent of an air parcel. A moist adiabatic ascent causes a given increase in surface temperature to be amplified with height. An early-generation general circulation model (GCM) that parameterized subgrid-scale convection by using dry and moist convection lapse rate adjustments, showed that an imposed 2 °C warming (cooling) of sea surface temperatures (SSTs) caused a +2.4 °C (–2.78 °C) response of simulated atmospheric temperatures at 1.5 km, and a +3.52 °C (–3.45 °C) response at 4.5 km (ref. 14). Such an amplification of surface warming with height was also found over the tropics in a classical $4 \times \text{CO}_2$ GCM experiment¹⁵ and was observed in the real atmosphere over the tropics¹⁶. A shift to a more moist adiabatic lapse rate in low latitudes as a result of an enhanced hydrological cycle has been suggested, which could influence the height of the freezing level surface¹⁷ or the height at which orographic clouds form.

The rapid melting of high mountain glaciers in low latitudes^{18,19} has been related to the observed warming of tropical ocean surfaces¹⁷, which in turn correlates with increases in tropospheric

Table 1 Simulated values of parameters (GENESIS GCM)

Parameter	Monteverde	Serrenia	Mt Kinabalu	Mt Virunga
1 $\Delta T(2 \times \text{CO}_2)$	2.1	2.1	2.2	2.5
2 $\Delta T(\text{LGM})$	–2.4	–2.2	–2.1	–1.9
3 $\Delta \text{AH}(2 \times \text{CO}_2)$	16	13	17	15
4 $\Delta \text{AH}(\text{LGM})$	–11	–15	–15	–13
5 $\Delta W(2 \times \text{CO}_2)$	25.3	25.0	25.9	29.4
6 $\Delta W(\text{LGM})$	–28.6	–26.8	–25.5	–22.9
7 $\Delta Z_{\text{RH}}(2 \times \text{CO}_2)$ DJF	228	109*	176	98
8 $\Delta Z_{\text{RH}}(2 \times \text{CO}_2)$ JJA	–127	–34	–20	–156
9 $\Delta Z_{\text{RH}}(\text{LGM})$ DJF	93	–40*	–454	–111
10 $\Delta Z_{\text{RH}}(\text{LGM})$ JJA	–103	270	625	–43.8
11 $\Delta Z_{\text{AH}}(2 \times \text{CO}_2)$ DJF	301	295	332	341
12 $\Delta Z_{\text{AH}}(2 \times \text{CO}_2)$ JJA	276	227	285	183
13 $\Delta Z_{\text{AH}}(\text{LGM})$ DJF	–162	–	–480	–242
14 $\Delta Z_{\text{AH}}(\text{LGM})$ JJA	–290	–168	–236	–161
15 $\Delta Z_{\text{W}}(2 \times \text{CO}_2)$	442	329	492	442
16 $\Delta Z_{\text{W}}(\text{LGM})$	–424	–298	–439	–320
17 LCL (DJF)	657	1,082	543	1,521
18 LCL (JJA)	510	719	577	1,634

Rows 1–6 show the difference in the annual temperature (°C), absolute humidity (AH; %) and warmth index (W ; °C-months) between the alternate climate experiments and the control run for each cloud forest location. The ΔZ rows (7–16) show the altitude shift (metres) required to reproduce the target proxy in the alternative climates (see text for details), averaged over summer (JJA) and winter (DJF) months, where appropriate. Rows 7–10 show the shift for RH; rows 11–14 for AH; and rows 15, 16 for W . A dash indicates that the proxy value was not reproduced in the alternative climate simulation for any of the 3 months in the season; an asterisk indicates that the proxy was reproduced in only one month. Rows 17, 18, seasonally averaged LCL.

moisture content^{16,20,21}. Specifically, analysis of decadal-scale trends of specific humidity, relative humidity, and temperature over the tropical western Pacific revealed that all have increased throughout the troposphere since the mid-1970s (ref. 16). GCMs driven by these observed increases in tropical SSTs reproduce the observed tropospheric warming from the 1970–1992 period primarily by hydrological cycle responses (that is, an increase in the global release of latent heat²⁰). Moreover, an intensifying tropical hydrological cycle and associated changes in tropical circulation are suggested by analysis of COADS data²². However, although lower atmospheric specific humidity and temperature are very likely to increase in response to oceanic surface warming, cloud formation depends on relative humidity, and thus the influence of surface warming on the probability of cloud formation at current cloud forest altitudes is not straightforward.

A direct estimate for the altitude of cloud formation is the point at which supercooled water condenses onto cloud condensation nuclei, corresponding to a vapour relative humidity of over 100%. This microscale description cannot be applied explicitly in GCMs, because they have horizontal grid resolutions larger in size than any individual cloud and vertical resolutions of several hundreds of metres. Parametric representations of such sub-grid-scale phenomena, like clouds, have been tested²³, and show that GCMs do predict large-scale (that is, grid-box-scale) variables which are highly correlated with cloud formation and occurrence (for example, water vapour, lapse rate, temperature and vertical velocity). We have explored the use of two quantities that are determined by these variables: the grid-box-averaged relative humidity (RH) and the LCL.

The simulation results for the LCL are presented in Table 1. Neither the GENESIS model nor another GCM from the University of Illinois (data not shown) produces grid-scale LCLs in the control simulation that coincide with current cloud forest altitudes. Indeed, GCMs are not expected to produce highly accurate simulations of the absolute value of hydrological parameters like the amount of water vapour at individual grid boxes²³, particularly near the steep topography that characterizes cloud forest locations. Thus, we focus on the RH as a grid-scale cloud formation proxy, especially differences in the RH between simulations.

Relative humidity changes are not easily predictable as it is not obvious in advance whether specific humidity increases enough locally in the $2 \times \text{CO}_2$ experiment, for example, to overcome the increase in saturation vapour pressure locally. In some GCM experiments with imposed SST increases¹⁴, RH decreased on average at an altitude of about 3 km (a typical cloud height) even though it increased near the surface. Similar results were obtained for a quadrupling of CO_2 (ref. 15). In Fig. 1b, our results show that the RH surface does not exhibit a consistent behaviour at the four cloud forest sites in the LGM simulation. However, the RH surface is consistently shifted upwards in the Northern Hemisphere winter season (December, January, February; DJF) at all of the sites in the $2 \times \text{CO}_2$ simulation. Opposite sign results for ΔZ_{RH} are obtained for the Northern Hemisphere summer season (June, July, August; JJA) at these same sites. We caution that the ΔZ_{RH} result should be viewed only as a crude proxy for cloud height change. Nevertheless, the RH surface of the Monteverde grid location (see Methods), for example, does suggest for the $2 \times \text{CO}_2$ case a rise of over 200 m in cloud height in winter, part of the dry season when that cloud forest relies most on the horizontal precipitation from cloud mists. Such a rise would imply a reduced frequency of contact and less horizontal precipitation, which would probably be of biological and hydrological significance to the cloud forest's structure and functioning.

As an alternative to predicting cloud heights to assess the impact of climate change on cloud forests, we relied upon the principles of biogeography models to predict the location of cloud forests in alternate climates. These models predict ecosystem locations using cutoffs in temperature sums and moisture balance calculations²⁴. One temperature sum is the warmth index (W_1), which represents

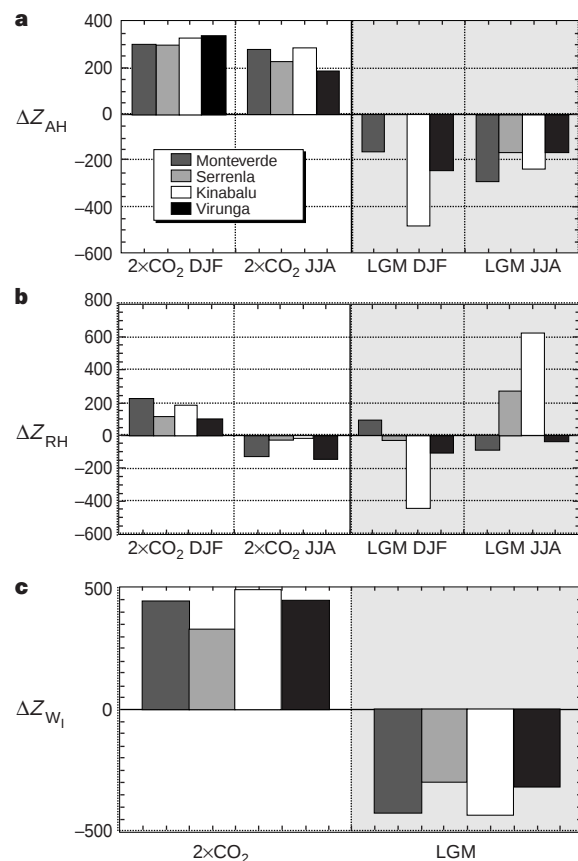


Figure 1 The shift in altitude, ΔZ , from the current cloud forest altitude that is required to reproduce a target climatic value (the variable's value at the cloud forest elevation in a GENESIS GCM control simulation) in a $2 \times \text{CO}_2$ or LGM simulation. The signal for summer (JJA) and winter (DJF) shifts are shown for **a**, the absolute humidity; and **b**, relative humidity; in **c**, only the yearly sum for the warmth index is shown. Left panels, ΔZ in the $2 \times \text{CO}_2$ simulation for four cloud forest locations; right panels, the LGM shift.

the sum of all monthly mean temperatures exceeding 5°C (ref. 1). The upper limit of the cloud forest on Mt Kinabalu (see Methods) corresponds to a cutoff in the W_1 of $85^\circ\text{C-months}^{25}$, perhaps a result of the linear relationship between the W_1 and potential evapotranspiration²⁶. Actual evapotranspiration varies linearly with net above-ground productivity²⁴, which can be successfully tied to ecosystem type²⁷. Therefore, it is not surprising that the W_1 is found to correlate broadly with forest type. In addition, we calculate the absolute humidity (AH) as a measure of the atmospheric water content at cloud forest elevations because moisture balance is also an important determinant of ecosystem type.

Figure 1a shows altitude shifts in the cloud forest absolute humidity surface in the GENESIS GCM for the $2 \times \text{CO}_2$ (ref. 28) and LGM²⁹ experiments: all four widely separated current cloud forest locations experience an increase in absolute humidity when global surface temperatures are warmer, as well as a decrease when it is cooler. This is anticipated, as warmer surface temperatures associated with $2 \times \text{CO}_2$ simulations cause increased evapotranspiration and thus the altitude at which some absolute humidity surface would occur in the control experiment is expected to rise (it rises by ~ 300 m) in the $2 \times \text{CO}_2$ simulation; likewise, a descent (by ~ 200 – 500 m) in the LGM experiment is not unexpected.

Figure 1c shows the shift in altitude required to reproduce the W_1 in alternate climates. As expected, the W_1 decreases in the LGM experiment and increases in the $2 \times \text{CO}_2$ case. The predicted altitude shift in W_1 is very similar to that of AH, suggesting that both the temperature and the moisture conditions of the current

cloud forest will be reproduced at this shifted altitude. Because the warmth index and absolute humidity together reflect components of many contemporary biogeography models, the agreement in their altitude shifts is noteworthy. The results are suggestive because of the uniformity and comparable magnitude of many of our multiple proxy results, the agreement with ice-age pollen shifts, and our technique's reliance on simulation differences as opposed to absolute values (differences in GCM variables are often thought to be more reliable than absolute values).

The results of Pounds *et al.*³⁰, which track invasions of premon-tane species into cloud forest habitat near Monteverde, Costa Rica, provide further evidence of the climate sensitivity of these ecosystems. Notably, they have rejected habitat destruction pressures at lower elevations as a cause of this upslope movement. Thus, this and other cloud forests may be experiencing the dual stresses of changing microclimates and invading species from lower elevations, driven in part by changes in the height of orographic cloud bank formation in the dry season and/or increased evapo-transpiration. In light of these observations and our modelling results, the impact of climate change on cloud forests needs further investigation. In addition to using more sophisticated biogeography models and higher-resolution climate-change scenarios from additional GCMs, future work should involve GCMs driven by the observed trend in tropical SSTs^{17,20} to examine present-day changes affecting these ecosystems. More realistic analyses at each of our TCMF sites will require transient GCM simulations combined with statistical down-scaling techniques²³ and/or coupled mesoscale regional models that better resolve the steep topographic conditions and local land use changes at most cloud forest sites. Also, in addition to focusing on changes in the vertical distribution of cloud decks, changes in the latitudinal distribution of circulation systems and trade wind belts and the implications for cloud forests should be investigated. However, the implications of even our crude proxy results are clear: climate change will probably affect the distribution of the potential locations for cloud forests and, if our analysis proves credible, it may force those situated near mountain tops out of existence. □

Methods

We investigated the impact of climate change on four cloud forest locations, which span four continents, with a variety of altitudes and ocean proximities. The first site is the Monteverde cloud forest (84.8° W, 10.3° N, 1,500 m) in Costa Rica, which we chose owing to the breadth of climate, vegetation and zoological research ongoing there. As this location is an ocean cell in our GCM, we used the adjacent eastward land cell to examine a potential response at Monteverde. In South America, we chose the cloud forest at Serrania de Macuira (71.5° W, 12.5° N, 865 m), which sits on a peninsula on the northern shore of Colombia at a relatively low altitude. Mount Kinabalu (116.3° E, 6.5° N, 2,500 m), our third site, is located on the island of Borneo. Our final site, Mount Virunga (29.5° E, 1.8° S, 3,000 m) in Africa, is home to the severely endangered mountain gorilla. This high-altitude cloud forest is land-locked, although it is close to Lake Victoria.

We used the global climate model GENESIS version 2 (ref. 28) to predict changes in the elevation of cloud forests for three simulations: today (control), the LGM²⁹, and a CO₂-doubled atmosphere²⁸. The simulations include a dynamic sea-ice model, a land-surface transfer scheme, a six-layer soil model and a 50-m slab mixed-layer ocean. The resolution of the simulations is spectral T31 (~3.75° longitude and latitude), with 18 vertical levels. The CO₂ concentration for the LGM was reduced by a specific percentage relative to the control that corresponds to the CO₂ lowering from the pre-industrial era to the LGM, insolation was set for the Earth orbit at 21 kyr, and ICE-4g ice sheets were prescribed. The 2 × CO₂ simulation was identical to the control run, except that the CO₂ concentration was doubled from 345 to 690 p.p.m.v. For each simulation, we used the monthly means, further averaged over ten years of each run, of the following variables: the pressure *p*, the specific humidity *q*, the air temperature *T*, and the elevation *Z*, of each grid level above mean sea level.

We determined the altitude shift of the relative humidity proxy, for example, as follows. We computed the simulated relative humidity surface, RH_{CF}, at the altitude of each of the TCMF sites, *Z*_{CF}, for today's climate (the control). We

then compute, by linear interpolation (logarithmic interpolation produces similar results) between vertical levels in the GCM, the altitude, *Z*'_{CF}, at which RH_{CF} is reproduced at each TCMF grid box for both LGM and 2 × CO₂ climate change experiments. The difference, $\Delta Z_{RH} = Z'_{CF} - Z_{CF}$, is the predicted altitude shift. These results are shown in Table 1, together with calculations of the changes in temperature, absolute humidity and *W*₁. Note that we do not search for the altitude at which the RH surface reaches 100% in this technique, because in the real atmosphere cloudiness often persists at grid-box-averaged RH values of much less than 100%, when only part of a grid box is cloud-covered. The same technique is used for ΔZ_{AH} and ΔZ_{W_1} . Comparison of the *W*₁ in climate change minus control simulations relies only on temperature differences: thus, major systematic errors in absolute temperature should cancel out. To estimate the error in our altitude shift calculations requires knowledge of the accuracy of the 2 × CO₂ simulations, which is unknown. We performed a simple error analysis by arbitrarily assuming that each monthly temperature in the 2 × CO₂ simulation was too warm by 1 °C: then the altitude shift error for the *W*₁ is 180 m, which is less than our calculated altitude shifts for this proxy.

Received 24 November 1998; accepted 25 February 1999.

1. Kira, T. *Terrestrial Ecosystems* (Kyoritsu, Tokyo, 1976).
2. Bruijnzeel, L. A. & Proctor, J. in *Tropical Montane Cloud Forests* (eds Hamilton, L. S., Juvik, J. O. & Scatena, F. N.) 38–78 (Springer, New York, 1995).
3. Stadtmüller, T. *Cloud Forests in the Humid Tropics: A Bibliographic Review* (The United Nations University, Tokyo, and Centro Agronomico Tropical de Investigación y Enseñanza, Turrialba, Costa Rica, 1987).
4. Bruijnzeel, L. A. *Hydrology of Moist Tropical Forests and Effects of Conversion: A State of Knowledge Review* (IHP-UNESCO Humid Tropical Programme, Paris, 1990).
5. Hamilton, L. S., Juvik, J. O. & Scatena, F. N. in *Tropical Montane Cloud Forests* (eds Hamilton, L. S., Juvik, J. O. & Scatena, F. N.) 1–23 (Springer, New York, 1995).
6. Leo, M. in *Tropical Montane Cloud Forests* (eds Hamilton, L. S., Juvik, J. O. & Scatena, F. N.) 198–211 (Springer, New York, 1995).
7. Colinvaux, P. A., De Oliveira, P. E., Moreno, J. E., Miller, M. C. & Bush, M. B. A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science* **275**, 85–88 (1996).
8. Stute, M. *et al.* Cooling of tropical Brazil (5 °C) during the last glacial maximum. *Science* **269**, 379–383 (1995).
9. Guilderson, T. P., Fairbanks, R. G. & Rubenstone, J. L. Tropical temperature variations since 20,000 years ago: modulating interhemispheric climate-change. *Science* **263**, 663–665 (1994).
10. Rind, D. & Petet, D. Terrestrial conditions at the last glacial maximum and CLIMAP sea-surface temperature estimates: Are they consistent? *Quat. Res.* **24**, 1–22 (1985).
11. Thompson, L. G. *et al.* Late glacial stage and Holocene tropical ice core records from Huascan, Peru. *Science* **269**, 46–50 (1995).
12. Thompson, L. G. *et al.* A 25,000-year tropical climate history from Bolivian ice cores. *Science* **282**, 1858–1864 (1998).
13. Barros, A. P. & Lettenmaier, D. P. Dynamic modeling of orographically induced precipitation. *Rev. Geophys.* **32**, 265–284 (1994).
14. Schneider, S. H., Washington, W. M. & Chervin, R. M. Cloudiness as a climatic feedback mechanism: effects on cloud amounts of prescribed global and regional surface temperature changes in the NCAR-GCM. *J. Atmos. Sci.* **35**, 2207–2221 (1978).
15. Manabe, S. & Wetherald, R. T. On the distribution of climate change resulting from an increase in CO₂ content of the atmosphere. *J. Atmos. Sci.* **37**, 99–118 (1980).
16. Gutzler, D. S. Climatic variability of temperature and humidity over the tropical western Pacific. *Geophys. Res. Lett.* **19**, 1595–1598 (1992).
17. Diaz, H. F. & Graham, N. E. Recent changes in tropical freezing heights and the role of sea surface temperature. *Nature* **383**, 152–155 (1996).
18. Hastenrath, S. & Kruss, P. The dramatic retreat of Mount Kenya's glaciers 1963–87: greenhouse forcing. *Ann. Glaciol.* **16**, 127–133 (1992).
19. Thompson, L. G. *et al.* Recent warming: ice core evidence from tropical ice cores with emphasis on central Asia. *Glob. Planet. Change* **7**, 145–156 (1993).
20. Graham, N. E. Simulation of recent global temperature trends. *Science* **267**, 666–671 (1995).
21. Gaffen, D. J., Barnett, T. P. & Elliott, W. P. Space and time scales of global tropospheric moisture. *J. Climate* **4**, 989–1008 (1991).
22. Flohn, H. & Kapala, A. Changes of tropical air-sea interaction processes over a 30-year period. *Nature* **338**, 244–246 (1989).
23. IPCC (Intergovernmental Panel on Climate Change) *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) (Cambridge University Press, Cambridge, 1996).
24. Woodward, F. I. *Climate and Plant Distribution* (Cambridge University Press, Cambridge, 1987).
25. Kitayama, K. An altitudinal transect study of the vegetation on Mount Kinabalu, Borneo. *Vegetatio* **102**, 149–171 (1992).
26. Fang, J.-Y. & Yoda, K. Climate and vegetation on China II. Distribution of main vegetation types and thermal climate. *Ecol. Res.* **4**, 71–83 (1989).
27. Neilson, R. P. A model for predicting continental-scale vegetation distribution and water balance. *Ecol. Appl.* **5**, 362–385 (1992).
28. Thompson, S. L. & Pollard, D. Greenland and Antarctic mass balances for present and doubled CO₂ from the GENESIS version-2 global climate model. *J. Climate* **10**, 871–900 (1997).
29. Pollard, D. & Thompson, S. L. Climate and ice-sheet mass balance at the last glacial maximum from the GENESIS version 2 global climate model. *Quat. Sci. Rev.* **16**, 841–864 (1997).
30. Pounds, J. A., Fogden, M. P. L. & Campbell, J. H. Biological response to climate change on a tropical mountain. *Nature* **398**, 611–615 (1999).

Acknowledgements. We thank J. A. Pounds for his Monteverde temperature and precipitation records and for sharing his observational data; J. A. Pounds, P. Colinvaux, H. Diaz, N. Graham, J. Harte, S. Hastenrath and J. Randerson for comments; D. Pollard for help with the GENESIS datasets (P.N.E.); T. Root for discussions (S.H.S.); and M. Schlesinger for the LCL data from the UI simulations. C.J.S. is supported by an EPA STAR graduate fellowship; S.H.S. acknowledges partial support from the Winslow Foundation.

Correspondence and requests for materials should be addressed to S.H.S. (e-mail: shs@leland.stanford.edu).

Biological response to climate change on a tropical mountain

J. Alan Pounds^{*†}, Michael P. L. Fogden^{*†}
& John H. Campbell^{*}

^{*} Golden Toad Laboratory for Conservation, Monteverde Cloud Forest Preserve and Tropical Science Center, Santa Elena, Puntarenas 5655, Box 73, Costa Rica, Central America

[†] Department of Biology, University of Miami, PO Box 249118, Coral Gables, Florida 33124-0421, USA

Recent warming has caused changes in species distribution and abundance^{1–3}, but the extent of the effects is unclear. Here we investigate whether such changes in highland forests at Monteverde, Costa Rica, are related to the increase in air temperatures that followed a step-like warming of tropical oceans in 1976 (refs 4, 5). Twenty of 50 species of anurans (frogs and toads) in a 30-km² study area, including the locally endemic golden toad (*Bufo perigrinus*), disappeared following synchronous population crashes in 1987 (refs 6–8). Our results indicate that these crashes probably belong to a constellation of demographic changes that have altered communities of birds, reptiles and amphibians in the area and are linked to recent warming. The changes are all associated with patterns of dry-season mist frequency, which is negatively correlated with sea surface temperatures in the equatorial Pacific and has declined dramatically since the mid-1970s. The biological and climatic patterns suggest that atmospheric warming has raised the average altitude at the base of the orographic cloud bank, as predicted by the lifting-cloud-base hypothesis^{9,10}.

This hypothesis builds on evidence that rising sea surface temperatures (SSTs) have altered the climates of tropical mountains. Enhanced evaporation from warm ocean surfaces has generated large amounts of water vapour, and latent heat released as this

moisture condenses has accelerated atmospheric warming⁵. Because vertical thermal profiles have tended towards a moist adiabatic lapse rate, the decline in temperature with increasing elevation has diminished, amplifying the warming in the highlands relative to the lowlands^{11–13}. Freezing heights have shifted upwards¹¹, and glaciers on high tropical mountains are rapidly melting¹⁴. If temperature-dependent relative humidity surfaces, and thus cloud-formation heights, have likewise shifted upwards¹⁰, organisms may be affected in various ways. Monteverde's stratus-stratocumulus bank, which forms as the trade winds meet the Caribbean slope of the Cordillera de Tilarán, flow upwards and cool adiabatically, influences several key ecological processes. A lifting cloud base should alter regional hydrology by reducing critical dry-season inputs of mist (low-intensity windblown precipitation) and cloud water (non-precipitating droplets deposited onto vegetation)^{15,16}.

To examine climate trends, we have analysed patterns of precipitation, stream flow, air temperatures and SSTs. The rainfall and air-temperature data (collected by J.H.C.) are from leeward cloud forest (1,540 m; ~1 km west of the Monteverde Cloud Forest Preserve headquarters and ~3 km west of the continental divide). The weather station lies on the western boundary of our 30-ha study plot for anoline lizards, which overlaps a 40-ha plot for birds. Both plots lie within the 30-km² anuran study area. The stream-flow data (from the Costa Rican Electrical Institute) are for the Río Cañas at Líbano (300 m; ~23 km northwest of Monteverde). The SST data (from NOAA) are for the Niño-3 region of the equatorial Pacific (~850 km SW; 150° W–90° W, 5° N–5° S).

Using daily rain-gauge records from the dry season (January–April), we develop a mist-frequency index. These data underestimate windblown inputs¹⁶ but contain signatures of mist events (as well as rare convective showers and wind-driven rains during cold fronts). Because 'dry days' (ones without measurable precipitation) represent intervals with little or no mist, their pattern conveys information about the incidence and temporal distribution of mist events. The frequency of dry days is inversely related to mist frequency, so we use the former as a negative index of the latter.

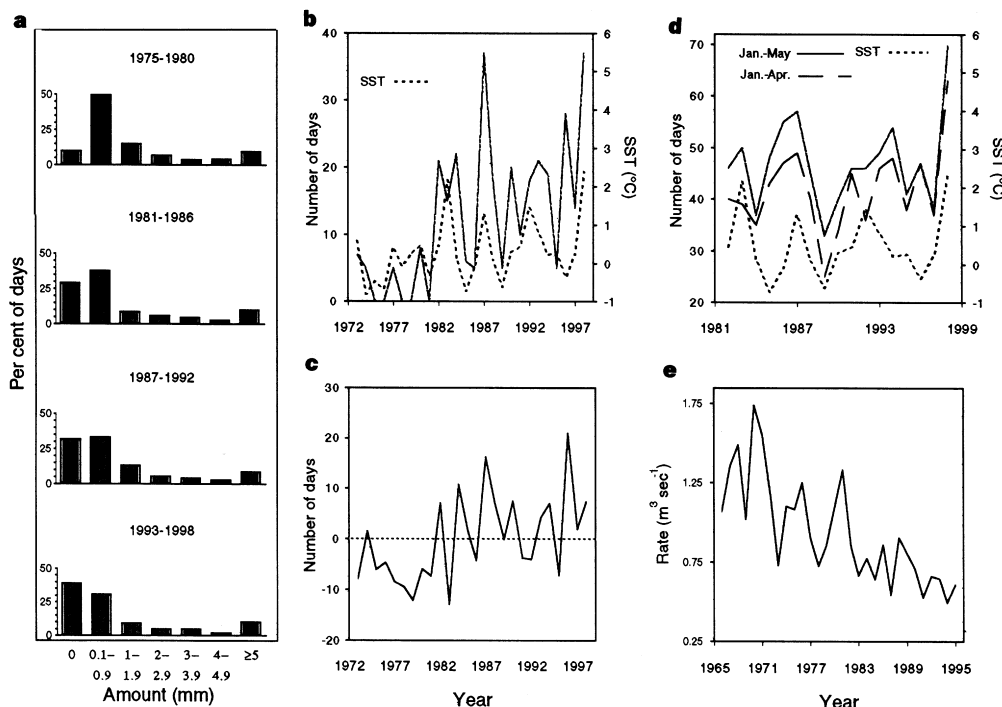


Figure 1 Trends and fluctuations in dry-season precipitation and stream flow. **a**, Frequency distributions of daily rainfall by 6-yr periods. **b**, SSTs (mean anomalies) and dry days in runs ≥ 5 . In the stepwise multiple regression, both SST and year explained significant amounts of variation in the number of dry days

(t -tests for partial coefficients, $P < 0.01$; for the full model, $F_{2,23} = 14.4$, $P < 0.001$; $r^2 = 0.56$). **c**, The drying trend, as illustrated by the residuals from step 1 of this regression (that is, the effect of year partialled on SST). **d**, Recent ENSOs and total dry days. **e**, Annual minimum stream flow (minimum daily average).

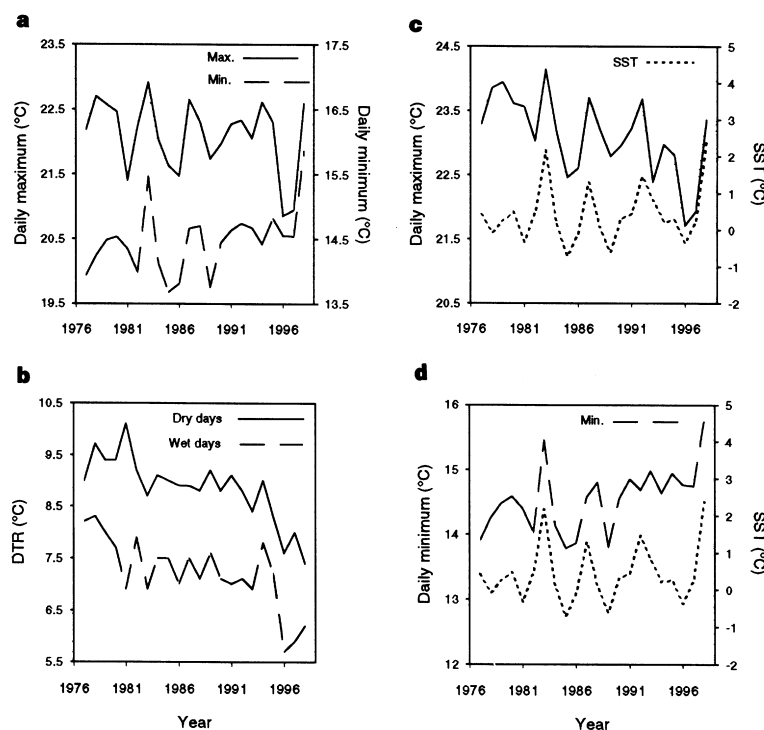


Figure 2 Trends and fluctuations in dry-season air temperatures. **a**, Daily minima and maxima (means). Test for trends: for the mean minimum, $r_s = 0.51$, $P < 0.016$, $N = 22$; for the mean maximum, $r_s = -0.19$, $P > 0.4$; for the mean DTR, $r_s = -0.64$, $P < 0.0002$. **b**, Diurnal temperature range (DTR) for wet days and

dry days (ones with and without measurable precipitation). Tests for trends: for wet days, $r_s = -0.64$, $P < 0.002$; for dry days, $r_s = -0.78$, $P < 0.0001$; **c**, SSTs (mean anomalies) and daily maxima on dry days. **d**, SSTs and daily minima on wet days.

As predicted by the lifting-cloud-base hypothesis, this index indicates a decline in mist frequency. Annual, seasonal and monthly rainfall totals show no trends over 1973–98. Day-to-day variability (the coefficient of variation of daily totals) likewise shows no trend for the wet season (May–October) or the transitional season (November–December). It exhibits a positive trend, however, for the dry season, mainly because dry days have increased in frequency (Fig. 1a). They have also increasingly coalesced into runs (Fig. 1b; up to 16 days). Although these dry periods are associated with warm episodes of the El Niño/Southern Oscillation (ENSO), the drying trend remains significant after ENSO-scale fluctuations are taken into account (Fig. 1b, c). Warm episodes, together with this underlying trend, produced the four largest recorded peaks in total dry days (in 1983, 1987, 1994 and 1998; Fig. 1d). Annual minimum stream flow, correlated with mist frequency, likewise indicates an ENSO-related signal superimposed on a drying trend (Fig. 1e).

Trends in dry-season air temperatures also accord with the lifting-cloud-base hypothesis. Because the daily minimum has increased relative to the daily maximum, the diurnal temperature range (DTR, the difference between the two) has declined (Fig. 2a). The negative trend implies increasing cloudiness, which lowers daytime temperatures by blocking solar radiation and raises night-time temperatures by reducing radiative heat losses^{17,18}. The coefficient of variation of DTR shows no trend, suggesting that day-to-day variability of cloudiness has not changed. Dry days exhibit larger DTRs than 'wet days' (ones with measurable precipitation), but the negative trend in DTR applies to both (Fig. 2b). The former result agrees with our observation that dry days often occur under clear skies. The trend for dry days implies that their numbers have increased by the addition of dry, cloudy days, as would be expected if a lifting cloud base has increased the fraction of non-precipitating clouds. This trend is not due entirely to changes in daily minima: daily maxima have declined, suggesting afternoon cloudiness on the added dry days (Fig. 2c). The nocturnal warming trend is strongest for wet days (Fig. 2d). Both daily maxima and minima are correlated

with SSTs (Fig. 2c, d).

The trends in DTR run counter to other proposed explanations for the decline in mist frequency: (1) the frequency of cloud-bank formation may have decreased; or (2) the average position of the cloud bank's leeward (western) edge may have receded eastward, beyond our weather station, if lee warming and drying have changed. These hypotheses, which envisage an increased frequency of clear skies over this station, predict a decrease in cloudiness or an increase in its day-to-day variability.

Changes in bird abundance at 1,540 m support the prediction that species should respond to climate change according to their distributions along climatic gradients¹⁹. Examining data for 1979–98 (collected by M.P.L.F.), we calculate trends in abundance (inferred from observation frequency) for two categories of species: (1) 'lower montane'—the breeding fauna of cloud forest ($\geq 1,470$ m) in the baseline years (1979–81); and (2) 'premontane'—cloud-forest-intolerant species, which bred only in premontane areas ($< 1,470$ m) during 1979–81. In the 1,540-m plot, the former shows no consistent pattern, whereas the latter chiefly exhibits increases (Fig. 3a).

Accordingly, the number of lower-montane species present in this plot has remained comparatively stable, whereas that of premontane species has increased (Fig. 3b). The increase has not been monotonic, because species that have extended their distributions upslope have often shifted back down. Net colonization (the number moving up minus the number moving down) is positive in some years and negative in others. The average rate, estimated as the slope of the linear regression, is 19 species per decade (excluding wide-ranging raptors and swifts, often recorded as flyovers). Colonizing species represent a broad taxonomic and ecological range and have come from premontane areas of both the Caribbean and Pacific slopes.

Fifteen colonizing species have established breeding populations in the 1,540-m plot, which have generally increased in density. For instance, golden-crowned warblers (*Basileuterus culicivorus*) first

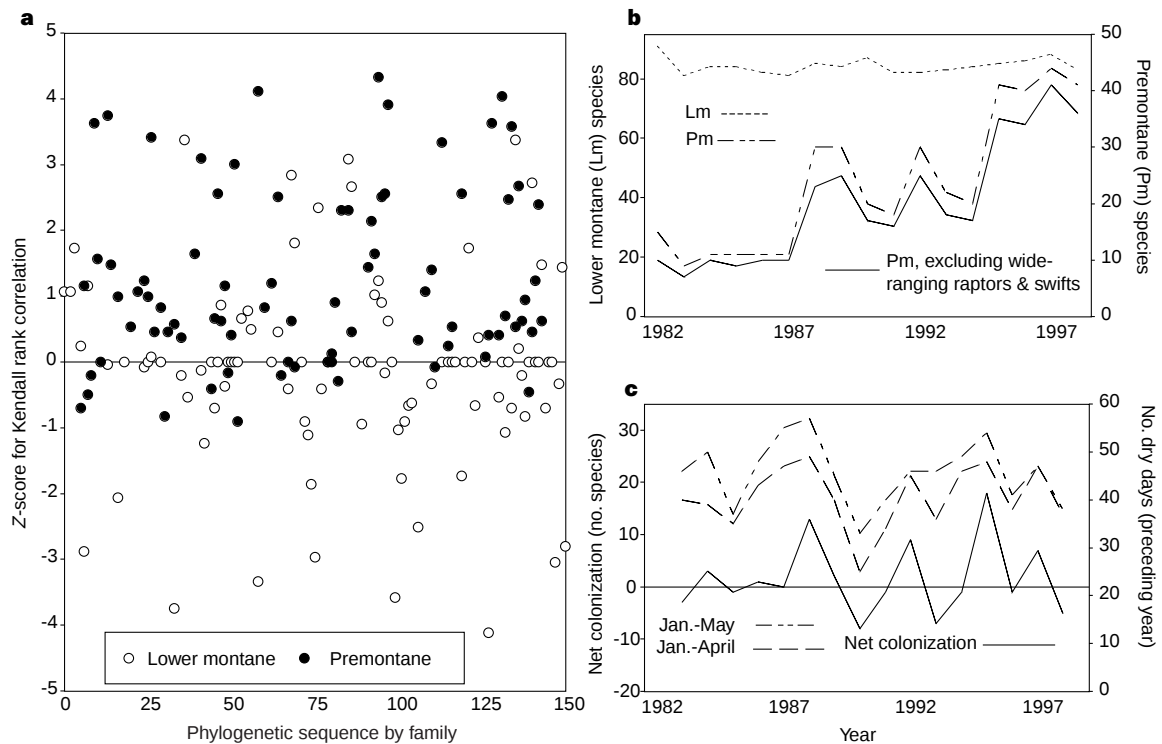


Figure 3 Changes in the bird community of the 1,540-m plot. **a**, Trends in abundance of individual species in relation to breeding distribution (lower montane versus premontane) in the baseline years. Values are normal approximations (Z) for Kendall's τ corrected for ties. Positive scores indicate increases, negative ones decreases. **b**, Trends in numbers of species present.

For lower-montane species, $\tau = 0.18$, $P > 0.34$, $N = 17$; for premontane species, $\tau = 0.67$, $P < 0.005$. **c**, Climate and upslope colonization. In the multiple regression, number of dry days in the preceding year's dry season accounted for 53% of the variation in net colonization.

nested there in 1994 (3 pairs) but increased to 20 pairs by 1998. Lesser greenlets (*Hylophilus decurtatus*) first nested there in 1988 (2 pairs) but had increased to 12 pairs by 1998. Although some recent arrivals remain at low density, their presence nevertheless represents significant change. Keel-billed toucans (*Ramphastos sulfuratus*; 2 pairs in 1998) previously bred only in lowlands and foothills²⁰; since 1995 they have nested alongside resplendent quetzals (*Pharomachrus moccino*), which symbolize Middle American cloud forests.

Habitat alteration has contributed to changes in distribution and abundance but cannot easily explain the patterns at 1,540 m. The mosaic of forest and clearings outside the Monteverde Preserve dates to the 1940s and early 1950s; no major deforestation has occurred in recent decades. This mosaic, which lies mostly in premontane areas, should favour upslope colonization by open-country birds, which increase in abundance relative to forest birds where there are clearings. The patterns at 1,540 m, however, apply to both.

Moreover, these patterns are strongly related to variation in mist frequency. Of various climatic variables tested by multiple regression, total dry days in the preceding year's dry season is by far the best predictor of upslope movements (Fig. 3c). A model based on 1982–96 patterns⁹ correctly predicted the sign and magnitude of net colonization in 1997 and 1998. Although a weak response to the 1982–83 El Niño is evident, the first major influx of colonists followed the 1986–87 warm episode.

Anoline lizard populations, which are sensitive to rainfall variability²¹, have declined in association with the same mist-frequency patterns. In the 1,540-m plot, two highland species endemic to Costa Rica and western Panama (the cloud forest anole, *Norops tropidolepis*, and the montane anole, *N. altae*) began to decline in the late 1980s and had disappeared by 1996 (Fig. 4a). In contrast to these species, the gray lichen anole (*N. intermedius*), which is most abundant in warmer, drier areas further down the Pacific slope^{8,22}, shows no trend at 1,540 m (ref. 8). For each

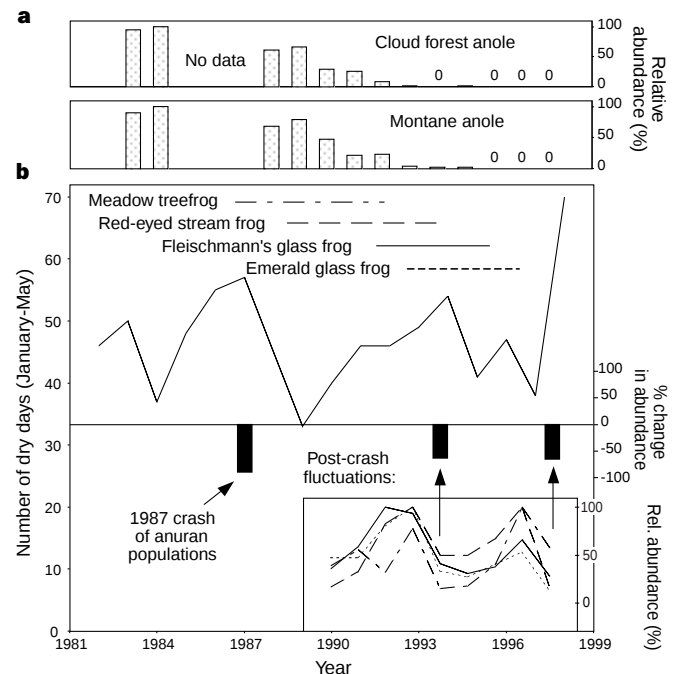


Figure 4 Declines in populations of anoline lizards and anurans. **a**, Anoles. For correlations of abundance with number of dry days in the preceding dry season, $\tau \leq -0.55$, $P < 0.016$, $N = 11$ yrs. **b**, Anurans. The random probability (P) that a population downturn would correspond to one of the 3 largest peaks in total dry days equals the proportion of years exhibiting these peaks. Thus, for 1984–98, $P = 3 \text{ peaks} \div 15 \text{ yr} = 0.2$ (a conservatively high estimate because the climatic extremes are the largest during 1973–98). Substituting 0.2 in a binomial distribution, we calculate the maximum probability that all three demographic events would, by chance alone, correspond to these extremes to be 0.008.

declining species, wet-season abundance is negatively correlated with total dry days in the preceding dry season but shows no correlation with other climatic variables tested.

The anuran declines, although more episodic than the anole declines, are likewise associated with mist-frequency patterns. To follow changes after the 1987 crashes, J.A.P. and co-workers have monitored selected populations since 1990⁷. Wet-season abundances of four aquatic-breeding species, which have fluctuated far below pre-crash levels, underwent synchronous downturns in 1994 and 1998 (Fig. 4b). Thus, three demographic events—in 1987, 1994, and 1998—correspond to the three largest peaks in total dry days. The probability that all three events would, by chance alone, correspond to these climatic extremes is ≤ 0.008 (Fig. 4b). Because anuran populations ordinarily fluctuate with climate, the 1987 crashes conceivably were due to an external event (such as the arrival of a pathogen) that coincided by chance with the 1986–87 warm episode. Similarities between demographic events suggest, however, that the climatic patterns are more than incidental to the declines. For example, red-eyed stream frogs (*Hyla uranochroa*) at several sites underwent extinctions in both 1987 and 1998, the most extreme years climatically. (The frogs had recolonized in the intervening years.) Populations of terrestrial-breeding rain frogs (*Eleutherodactylus diastema*) likewise crashed over large areas in 1987 and 1998 (J.A.P., unpublished results).

The changes in populations of birds, lizards and anurans are concordant yet diverse. All are associated statistically with the same climatic patterns and occurred simultaneously, implying a broad response to regional climate change which crossed an important threshold in the late 1980s. Although one study concluded that anuran declines in Australia are probably not climate-related²³, the analyses relied on monthly weather data. The reduction in mist frequency reported here, which agrees with large-scale climate trends^{5,11–13} and simulations of greenhouse warming¹⁰, is evident only in the daily records. The diversity of demographic changes related to this reduction suggests that climate has orchestrated them by means of several biotic mechanisms (although climate may also interact with other physical factors⁶). Evidence that dry weather in 1983 increased the vulnerability of harlequin frogs (*Atelopus varius*) to lethal parasites along one stream^{24–26} inspired the “climate-linked epidemic hypothesis”²⁶. As the habitat dried and the frogs gathered near waterfalls, their probability of being attacked by parasitic flies increased sharply: forty dead or dying frogs were observed. Because climate affects host–parasite relationships²⁷ and amphibians^{28,29} in various ways, it may have set the stage for similar mortality events, including those ascribed to chytrid fungus outbreaks³⁰. If so, the recent, widespread amphibian extinctions in seemingly undisturbed highland forests may attest to how profound and unpredictable the outcome can be when climate change alters ecological interactions. □

Methods

Precipitation and stream flow. We examine rainfall trends for 1973–98 ($N = 26$; Spearman rank correlations, r_s ; $\alpha = 0.05$; 2-tailed). Time series are short considering potential interdecadal variability but cover the critical period of oceanic warming. Analyses of dry-season frequency distributions (Fig. 1a) focus on January–April, but we tally dry days (Fig. 1b–d) for January–May. The dry season sometimes extends beyond April, depending on when convective rains begin. The number of days comprising dry periods has increased, regardless of the definition of minimum period length (≥ 1 , ≥ 2 , ..., ≥ 10 days). Whereas total dry days exhibits lag-1 autocorrelation, the number in runs ≥ 5 does not, and thus meets independence assumptions for F statistics. We regress the latter variable on SST and year in a stepwise analysis. SSTs are anomalies (deviations from the running mean) averaged for January–May. We screened SSTs from the different Niño regions (and the Southern Oscillation Index): Niño-3 values exhibit the strongest correlations with precipitation variables. To examine stream-flow patterns, we use Fourier analysis. Deforestation may have contributed to the downward trend by altering runoff patterns. The clearings,

however, date to the 1960s or earlier, whereas minimum flow has continued to decline in the 1980s and 1990s.

Air temperatures. We examine trends for 1977–98 ($N = 22$; Fig. 2). Because we use means for January–May to match our tallies of dry days, the inferred increase in cloudiness may involve convective cumulus and cumulonimbus, as well as advective status and stratocumulus. (Convection increases in frequency over January–May as the trade winds subside; it typically becomes the dominant pattern in early-to-middle May.) The negative trend in DTR, however, applies to data subsets that progressively reduce the influence of convection (Jan–Apr, Jan–Mar, Jan–Feb, and January), suggesting effects of advective clouds.

Birds. The 1979–81 survey of the Monteverde region included 13.5 months of sampling ($>2,000$ h). To calculate individual species trends (Fig. 3a), we examine observation frequencies (proportions of 6-day periods in which a species was present) over 1982–98 ($N = 17$). Number of periods sampled was ≥ 10 per primary breeding season (March–July) and similar for the remaining months (total, 499; $>15,000$ h). Sampling included daily listening to the dawn chorus from key points, plus surveys of a 2-km circuit (2 per period) within the 40-ha plot (25 ha of primary forest, 13 ha of secondary forest, 2 ha meadows). We examine observation frequencies weighted to equalize contributions of breeding and non-breeding seasons. Analyses of unweighted data, however, give virtually identical results, as do analyses for the breeding season only. To estimate changes in breeding density of recent arrivals, we compare observations in the year of first breeding to singing-male counts in 1998. Independent variables in the multiple regression analysis (Fig. 3c) include annual and seasonal mean daily rainfall, number of dry days (Jan–Apr, Jan–May, runs ≥ 3 , runs ≥ 5 , and longest run), and mean temperature (daily maximum, daily minimum, and DTR). We also include each variable with its effect lagged by 1 year.

Anoline lizards and anurans. We calculate the relative abundance of anoles (per cent of the maximum yearly average for 1983–98) from counts of adults during the wet season (≥ 10 2-h surveys per yr)⁸. Correlations with climatic variables are based on data through 1996, when both species had disappeared from the study area ($N = 11$). Variables tested are listed above (see Birds). Pounds *et al.*⁷ have described study sites and methods used to monitor meadow treefrogs (*Hyla pseudopuma*), red-eyed stream frogs (*Hyla uranochroa*), Fleischmann's glass frogs (*Hyalinobatrachium fleischmanni*), and emerald glass frogs (*Centrolenella prosoblepon*). For each species in each wet season, we average abundances of the different populations and calculate relative abundance (per cent of the maximum yearly average for 1990–98). Per cent change for 1987 is a conservative estimate derived from baseline observations and 1990 data for the same four species.

Received 24 November 1998; accepted 25 February 1999.

1. Parmesan, C. Climate and species' range. *Nature* **382**, 765–766 (1996).
2. Veit, R. R., McGowan, J. A., Ainley, D. G., Wahls, T. R. & Pyle, P. Apex marine predator declines 90% in association with changing oceanic climate. *Global Change Biol.* **3**, 23–28 (1997).
3. McGowan, J. A., Cayan, D. R. & Dorman, L. M. Climate–ocean variability and ecosystem response in the northeast Pacific. *Science* **281**, 210–217 (1998).
4. Trenberth, K. E. & Hurrell, J. W. Decadal atmosphere–ocean variations in the Pacific. *Clim. Dyn.* **9**, 303–319 (1994).
5. Graham, N. E. Simulation of recent global temperature trends. *Science* **267**, 666–671 (1995).
6. Pounds, J. A. & Crump, M. L. Amphibian declines and climate disturbance: the case of the golden toad and the harlequin frog. *Cons. Biol.* **8**, 72–85 (1994).
7. Pounds, J. A., Fogden, M. P. L., Savage, J. M. & Gorman, G. C. Tests of null models for amphibian declines on a tropical mountain. *Cons. Biol.* **11**, 1307–1322 (1997).
8. Pounds, J. A. in *Monteverde: Ecology and Conservation of a Tropical Cloud Forest* (eds Nadkarni, N. M. & Wheelwright, N. T.) (Oxford University Press, New York, in press).
9. Pounds, J. A., Fogden, M. P. L. & Campbell, J. H. in *Meeting Report, Bird Life International/WWF Workshop on the Impacts of Climate Change on Flora and Fauna* (ed. Briggs, B.) (Royal Society for Protection of Birds, Bedfordshire, UK, 1997).
10. Still, C. J., Foster, P. N. & Schneider, S. H. Simulating the effects of climate change on tropical montane cloud forests. *Nature* **398**, 608–610 (1999).
11. Diaz, H. F. & Graham, N. E. Recent changes in tropical freezing heights and the role of sea-surface temperature. *Nature* **383**, 152–155 (1996).
12. Beniston, M., Diaz, H. F. & Bradley, R. S. Climatic change at high elevation sites: an overview. *Clim. Change* **36**, 233–251 (1997).
13. Diaz, H. F. & Bradley, R. S. Temperature variations during the last century at high elevation sites. *Clim. Change* **36**, 253–279 (1997).
14. Thompson, L. G. *et al.* Late glacial stage and holocene tropical ice core records from Huascarán, Peru. *Science* **269**, 46–50 (1995).
15. Cavalier, J., Solis, D. & Jaramillo, M. A. Fog interception in montane forests across the Central Cordillera of Panamá. *J. Trop. Ecol.* **12**, 357–369 (1996).
16. Clark, K. L., Nadkarni, N. M., Schaefer, D. & Gholz, H. Atmospheric deposition and net retention of ions by the canopy in a tropical montane forest, Monteverde, Costa Rica. *J. Trop. Ecol.* **14**, 27–45 (1998).
17. Karl, T. R. *et al.* A new perspective on recent global warming: asymmetric trends of daily maximum and minimum temperatures. *Bull. Am. Meteorol. Soc.* **74**, 1007–1023 (1993).

18. Gómez, I. E. & Fernandez, W. Variación interanual de la temperatura en Costa Rica. *Top. Meteor. Oceanogr.* **3**, 27–44 (1996).
19. Karr, J. R. & Freeman, K. E. Habitat selection and environmental gradients: dynamics in the 'stable' tropics. *Ecology* **64**, 1481–1494 (1983).
20. Fogden, M. P. L. *An Annotated Checklist of the Birds of Monteverde and Peñas Blancas* (Green Mountain, Monteverde, Costa Rica, 1993).
21. Andrews, R. M. Population stability of a tropical lizard. *Ecology* **72**, 1204–1217 (1991).
22. Pounds, J. A. Ecomorphology, locomotion, and microhabitat structure: patterns in a tropical mainland *Anolis* community. *Ecol. Monogr.* **58**, 299–320 (1988).
23. Laurance, W. F. Catastrophic declines of Australian rainforest frogs: is unusual weather responsible? *Biol. Cons.* **77**, 203–212 (1996).
24. Crump, M. L. & Pounds, J. A. Lethal parasitism of an aposematic anuran (*Atelopus varius*) by *Notochaeta bufonivora* (Diptera: Sarcophagidae). *J. Parasitol.* **71**, 588–591 (1985).
25. Pounds, J. A. & Crump, M. L. Harlequin frogs along a tropical montane stream: aggregation and the risk of predation by frog-eating flies. *Biotropica* **19**, 306–309 (1987).
26. Crump, M. L. & Pounds, J. A. Temporal variation in the dispersion of a tropical anuran. *Copeia* **1989**, 209–211 (1989).
27. Dobson, A. & Carper, R. in *Global Warming and Biological Diversity* (eds Peters, R. L. & Lovejoy, T. E.) 201–217 (Yale Univ. Press, New Haven, Connecticut, 1992).
28. Beebe, T. J. C. Amphibian breeding and climate. *Nature* **374**, 219–220 (1995).
29. Stewart, M. M. Climate driven population fluctuations in rain forest frogs. *J. Herpetology* **29**, 437–446 (1995).
30. Berger, L. et al. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proc. Natl Acad. Sci.* **95**, 9031–9036 (1998).

Acknowledgements. We thank C. J. Still, P. N. Foster and S. H. Schneider for sharing their GCM simulations and for their comments on the manuscript; we also thank M. Brenes, R. W. Carlson, P. M. Fogden, G. R. Graves, J. E. Jackson, F. J. Joyce, K. L. Masters, J. M. Savage and N. T. Wheelwright for comments that improved the final version. J.A.P. acknowledges partial support from the MacArthur Foundation, Stanford University's Center for Conservation Biology, the Brookfield Zoo, the University of Miami, the US NSF, the Jessie Smith Noyes Foundation, the Organization for Tropical Studies, the Tropical Science Center and the University of Florida.

Correspondence and requests for materials should be addressed to J.A.P. (e-mail: goldtoad@sol.racs.co.cr).

Humans use internal models to estimate gravity and linear acceleration

Daniel M. Merfeld[†]*, Lionel Zupan^{*} & Robert J. Peterka

Neurological Sciences Institute, Oregon Health Sciences University,
1120 NW 20th Avenue, Portland, Oregon 97209, USA

* These authors contributed equally to this work.

Because sensory systems often provide ambiguous information, neural processes must exist to resolve these ambiguities. It is likely that similar neural processes are used by different sensory systems. For example, many tasks require neural processing to distinguish linear acceleration from gravity¹, but Einstein's equivalence principle states that all linear accelerometers must measure both linear acceleration and gravity. Here we investigate whether the brain uses internal models, defined as neural systems that mimic physical principles, to help estimate linear acceleration and gravity^{2–4}. Internal models may be used in motor control^{5–7}, sensorimotor integration^{8–10} and sensory processing^{11–14}, but direct experimental evidence for such models is limited. To determine how humans process ambiguous gravity and linear acceleration cues, subjects were tilted after being rotated at a constant velocity about an Earth-vertical axis. We show that the eye movements evoked by this post-rotational tilt include a response component that compensates for the estimated linear acceleration even when no actual linear acceleration occurs. These measured responses are consistent with our internal model predictions that the nervous system can develop a non-zero estimate of linear acceleration even when no true linear acceleration is present.

When present, inter-aural linear acceleration has been shown^{15–17} to elicit a compensatory horizontal eye movement, the vestibulo-ocular reflex (VOR). A similar horizontal VOR might occur if the nervous system uses an internal model to mimic the external physical relationship¹ (Fig. 1, $F = G - A$) between gravity (G),

linear acceleration (A), and specific gravito-inertial force (GIF, F). Specifically, we propose that an internal (neural) model of this physical relationship exists (Fig. 2, $\hat{F} = \hat{G} - \hat{A}$), with the difference between the internal estimates of gravity ($\hat{G}$) and linear acceleration ($\hat{A}$) equalling the estimated GIF ($\hat{F}$). A non-zero estimate of linear acceleration will arise whenever there is a mismatch between the internal estimate of gravity and the physiological measurement of GIF, and a compensatory horizontal VOR will be evoked whenever the estimate of linear acceleration includes a non-zero, inter-aural component ($\hat{A}_y$, Fig. 2c).

To investigate these predictions, eight subjects (including four experienced psychophysical observers) were seated upright and rotated at a constant angular velocity of 200° s^{-1} about an Earth-vertical axis for 150 s. The rotating subjects were then rapidly decelerated (100° s^{-2}) to a stop. Immediately after stopping, subjects were tilted by 90° about an axis through the centre of the head (at ear level) in 1.5 s, to one of eight evenly spaced orientations: nose-up (NU), nose-down (ND), right-down (RD), left-down (LD), and each of the four orientations midway between these principal orientations (NU-RD, NU-LD, ND-RD, ND-LD). The trial order was randomized and secret; the direction of rotation and the direction of tilt were counter-balanced. Binocular eye movements were recorded in the dark throughout motion and for 150 s following deceleration with bitebar-stabilized video cameras. The video system was calibrated using 29 fixation targets. Video data were analysed off-line to give horizontal eye position, which was digitally differentiated to give horizontal eye velocity. Fast phases were removed using a computer algorithm, with some manual editing by experienced personnel, leaving the slow phase velocity.

Constant velocity rotation followed by tilt ('post-rotational tilt') was chosen to elicit a mismatch between estimated gravity and true gravity (Fig. 2). Deceleration to a stop following constant velocity rotation induces a strong cue from the semicircular canals which indicates rotation, due to the peripheral dynamics of the semicircular canals¹⁸. This cue falsely signals angular velocity in the direction opposite to the preceding rotation, even though the subject is at rest. At the same time, graviceptors (for example, vestibular otolith organs) measure a constant, unchanging, gravito-inertial force, resulting in a contradictory sensory situation. The illusory sensations induced by this and similar paradigms are "confusing, unpleasant, and difficult to describe"¹⁹. Rotational cues are known to influence the perceived orientation of gravity²⁰, often leading to illusory tilt²¹. Therefore, we suggest that the yaw rotational cue following post-rotational tilt leads to illusory tilt and that the direction of the illusory tilt matches the direction of tilt that would occur if the rotational cue were accurate.

To measure illusory tilt, the four subjects who were experienced psychophysical observers verbally answered standard questions regarding their subjective tilt more than 2 min after tilting. (A

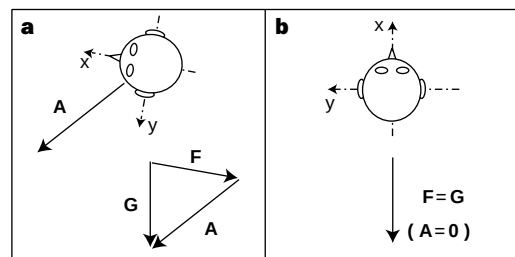


Figure 1 During acceleration on Earth, gravity (G) minus linear acceleration (A) yields specific gravito-inertial force (GIF, F). Different combinations of gravity and linear acceleration can yield the same measured GIF. **a**, GIF equals gravity minus linear acceleration ($F = G - A$ or $F + A = G$). **b**, GIF equals gravity alone. Since the measurement of GIF relative to the head is identical for these two situations, additional information is required for the nervous system to determine how much of the measured GIF is due to gravity and how much is due to linear acceleration.

[†] Present address: Vestibular Physiology Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114, USA.

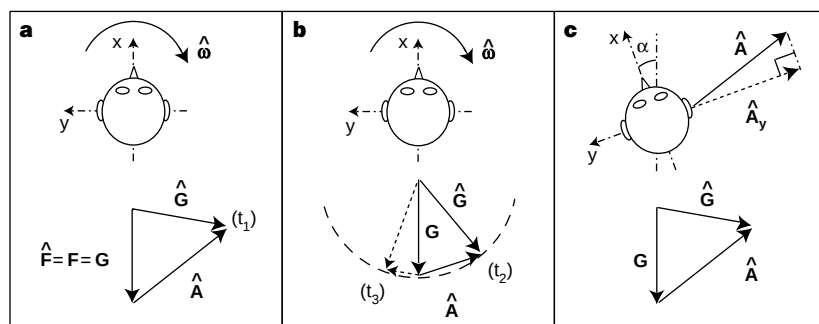


Figure 2 If the subject is oriented 'nose-up', a clockwise yaw rotational cue ($\hat{\omega}$) rotates the estimated gravity ($\hat{\mathbf{G}}$) toward the subject's right. **a**, When the estimated gravity is tilted relative to the actual orientation of gravity, the internal model will develop an estimated linear acceleration ($\hat{\mathbf{A}}$) that equals the estimated gravity ($\hat{\mathbf{G}}$) minus estimated GIF ($\hat{\mathbf{F}}$). This 'GIF resolution hypothesis' ($\hat{\mathbf{F}} = \hat{\mathbf{G}} - \hat{\mathbf{A}} \Rightarrow \hat{\mathbf{A}} = \hat{\mathbf{G}} - \hat{\mathbf{F}}$) is a neural representation of a physical relationship (that is, an internal model).

The estimated GIF ($\hat{\mathbf{F}}$) is assumed to equal the measurement of GIF, which, in this case, equals gravity ($\hat{\mathbf{F}} = \mathbf{F} = \mathbf{G}$). **b**, The estimates of gravity and linear acceleration are time varying. As the estimated tilt decreases (t_2) and even reverses (t_3), the estimated linear acceleration also decreases (t_2) and reverses (t_3). **c**, The projection of the estimated linear acceleration on the inter-aural axis ($\hat{\mathbf{A}}_y$) varies sinusoidally with subject orientation (α).

bitebar prohibited communication during the trial.) These subjects reported maximal illusory tilt in the predicted direction for 43 trials, no illusory tilt for 15 trials, and maximal illusory tilt opposite to the predicted direction for only 6 trials. Further, the subjects reported maximal illusory tilt in the predicted direction for 5 out of 5 trials when the illusory tilt exceeded 60° , and for 24 out of 26 trials when the illusory tilt exceeded 20° . Motion sickness symptoms, complex contradictory sensations and the delay (>2 min) between experiencing and reporting tilt probably contributed to reporting errors.

As subjective gravity did not match true gravity, the GIF resolution hypothesis (Fig. 2) predicts a non-zero estimate of linear acceleration. The non-zero projection of the estimated linear acceleration on the inter-aural axis should induce a horizontal VOR. This 'induced VOR' is similar to the VOR evoked by actual linear acceleration^{15–17} but is evoked by estimated linear acceleration even in the absence of actual linear acceleration. The induced VOR depends on rotation direction and subject orientation after tilt. The rotation direction influences the direction of illusory tilt, and subject orientation alters the projection of the estimated linear acceleration on the inter-aural axis (Fig. 2c). For example, the post-rotational response ($\hat{\omega}$) for a counterclockwise (CCW) rotation is clockwise (CW). For the nose-up subject orientation, estimated gravity ($\hat{\mathbf{G}}$) is tilted towards the subject's right, inducing an inter-aural component of estimated acceleration ($\hat{\mathbf{A}}_y$) toward the subject's right (Fig. 3a, upper sketch). This estimated acceleration induces a VOR to the left, which adds to the angular VOR. For the nose-down orientation following CCW rotation, the predicted tilt and inter-aural component of acceleration are towards the left, inducing a VOR to the right, which deducts from the angular VOR. Consistent with these predictions and with previous data²², the magnitude of the VOR following nose-up tilt was greater than that of the VOR following nose-down tilt after identical CCW rotations (Fig. 3a, upper traces). For the post-rotational responses following CW rotation, the predicted angular VOR and induced VOR are both reversed. Again, the induced VOR adds to the angular VOR for the nose-up orientation, but subtracts from it for the nose-down orientation. Consistent with these predictions, the magnitude of the horizontal post-rotational VOR for the nose-up orientation was greater than that for the nose-down orientation following identical CW rotations (Fig. 3a, lower traces).

For the left-down orientation, the predicted inter-aural component of estimated acceleration ($\hat{\mathbf{A}}_y$) is always towards the subject's right (Fig. 3b, sketch), inducing a VOR towards the left. This induced VOR should add to the angular VOR following CCW rotation and subtract from it following CW rotation. For the right-down orientation, the predicted induced VOR is towards the

right. Therefore, the induced VOR should add to the angular VOR following CW rotation and subtract from it following CCW rotation. As predicted, the left-down response magnitude was greater than the right-down response magnitude following identical CCW rotations and less than the right-down response magnitude following identical CW rotations (Fig. 3b).

For orientations midway between the primary tilt orientations, the comparisons are more complicated and also more noteworthy. Figure 3c shows the responses during NU-RD and ND-LD trials. Little difference was observed following CCW rotation, but a large difference was evident for the same orientations following CW rotation. In contrast, the responses during NU-LD and ND-RD trials show little difference following CW rotation but a large difference following CCW rotation (Fig. 3d). This pattern is consistent with the predictions of the GIF resolution hypothesis. For example, if the estimated linear acceleration is tilted approximately 45° to the right of vertical (Fig. 3c and d, CCW trials), the inter-aural component of this acceleration will be near-zero for one set of trials (Fig. 3c, CCW trials) and maximal for a second set of trials (Fig. 3d, CCW trials). For a similar tilt of estimated linear acceleration in the opposite direction, this pattern will reverse; the maximal inter-aural estimates of acceleration occur for those orientations previously showing near-zero estimates of acceleration (Fig. 3c, CW trials) and near-zero estimates of acceleration for those orientations previously showing maximal inter-aural linear acceleration (Fig. 3d, CW trials).

As shown in Fig. 3 (sketches), subject orientations separated by 180° (for example, nose-up and nose-down) result in opposite projections of the estimated linear acceleration on the inter-aural axis ($\hat{\mathbf{A}}_y$), leading to the prediction that the resulting induced VORs will be oppositely directed as well. Therefore, the induced VOR can be computed by subtracting the responses from subject orientations separated by 180° (for example, nose-up minus nose-down) and dividing the result by two. In performing this calculation, we assume that the underlying angular response is the same for these two orientations, since the rotational cue from the semicircular canals is identical. This assumption is consistent with studies showing that the duration of rotation sensation does not depend on subject orientation following 90° tilts²².

Figures 3e–h show the induced VOR for various orientations. It reverses after about 30 s for all orientations except for left-down and right-down. This reversal pattern provides further evidence of an induced VOR. The rotational cue reverses approximately 30 s after deceleration following constant velocity stimulation²³. Under the GIF resolution hypothesis, this canal reversal should cause small reversals in the estimate of tilt and linear acceleration, leading to a predicted reversal of the induced VOR for most subject orientations.

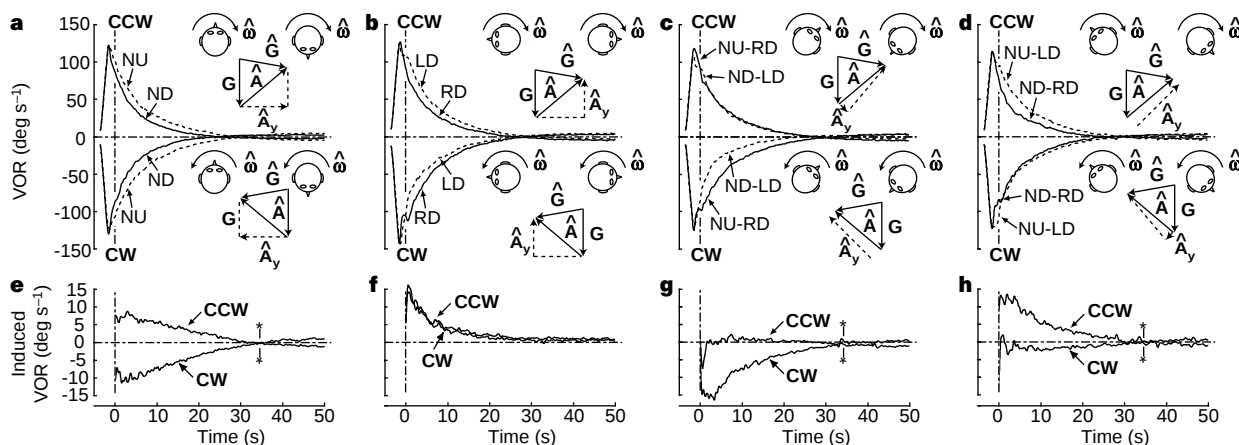


Figure 3 Post-rotational horizontal VOR. We label the post-rotational responses using the actual direction of rotation preceding the tilt (CW or CCW), but also sketch the oppositely directed post-rotational response ($\hat{\omega}$). Responses following CCW rotation are always positive (slow phase left); responses following CW rotation are always negative (slow phase right). Time 'zero' indicates completion of tilt. The post-rotational responses are shown for the following orientations:

a, Nose-up and nose-down, **b**, Left-down and right-down, **c**, NU-RD and ND-LD, and **d**, NU-LD and ND-RD. The induced VOR is calculated as: **e**, The nose-up response minus the nose-down response (divided by 2); **f**, The left-down response minus the right-down response (divided by 2); **g**, The NU-RD response minus the ND-LD response (divided by 2); and **h**, The NU-LD response minus the ND-RD response (divided by 2). Asterisks indicate induced VOR reversal.

However, in the right-down and left-down positions (Fig. 3f), the projection of the estimated linear acceleration on the inter-aural axis is always in a single direction. For example, in the left-down orientation (Fig. 3b), this projection is always towards the right following both CW and CCW rotations. Therefore, for the right-down and left-down orientations, reversal in the estimated tilt does not reverse the projection of the estimate of linear acceleration on the inter-aural axis, nor does it reverse the induced VOR.

The induced VOR and the sinusoidal fit to these data are plotted versus subject orientation (Fig. 4a) at one specific instant in time (3 s after tilt). The CW and CCW responses are well fit by sinusoids phase-shifted with respect to one another. The negative of the predicted estimate of inter-aural linear acceleration is shown versus subject orientation (Fig. 4b) for 45° tilts of the estimated linear acceleration. These predictions closely match the data (Fig. 4a).

Previous studies^{19,22,24} have shown that post-rotational tilts induce a response attenuation ('dumping') when graviceptor cues are

inconsistent with rotation cues. Our data confirm these findings, with the decay of the VOR always being more rapid with tilt than without tilt (not shown). Our data also show that the post-rotational responses vary systematically with head orientation. One explanation might be an orientation-dependent response attenuation. We reject this explanation for several reasons. First, if it were correct, subjective sensations of head rotation should similarly depend on head orientation. However, reports show that the duration of rotation sensation does not depend on subject orientation²². This was the case even when large, significant ($P < 0.01$) VOR time constant differences were measured during the same trials. Furthermore, as evident in Fig. 3b (CW, RD trace) and 3c (CW, NU-RD trace), the VOR actually increases immediately after the tilt in some orientations; this increase is inconsistent with any type of response attenuation but is consistent with the additive influence of an induced VOR. Finally, the decay rate caused by simple, orientation-dependent response attenuation should depend only on head orientation and should be independent of rotation direction. This attenuation is inconsistent with the experimentally observed dependence on rotation direction (Figs 3b–d and 4a).

We also considered and rejected other explanations for the eye movements. A static positional response was ruled out by measuring eye-movement responses after simply tilting each of the subjects by 90°; VOR responses were not observed after the tilt. An axis shift contaminant was rejected because vertical and torsional responses are small or non-existent in humans²⁵ following 90° post-rotational tilt. Our measurements of vertical and torsional responses (not shown) confirm these earlier findings. Three-dimensional calculations show that any horizontal VOR contaminant would be more than an order of magnitude smaller (1° s^{-1} maximum) than the measured induced VOR. A contaminant due to the dynamic 90° tilt was ruled out by measuring post-rotational responses following constant velocity (200° s^{-1}) yaw rotation about an Earth-horizontal axis ('barbecue-spit rotation') in four subjects. Subjects were stopped in the nose-up or nose-down orientation. The nose-up responses were always greater than the nose-down responses; the peak difference was approximately 10° s^{-1} for both rotation directions. These results are similar to the previously shown post-rotational tilt data (Fig. 3a, e), demonstrating that the dynamic 90° tilt is not responsible for the induced VOR.

Earlier studies have shown that information from the semicircular canals influences the interpretation of graviceptor cues^{3,20}; this influence is therefore included in mathematical models^{2,4,12,26}. A modelling simulation²⁷ for a 45° post-rotational tilt similar to

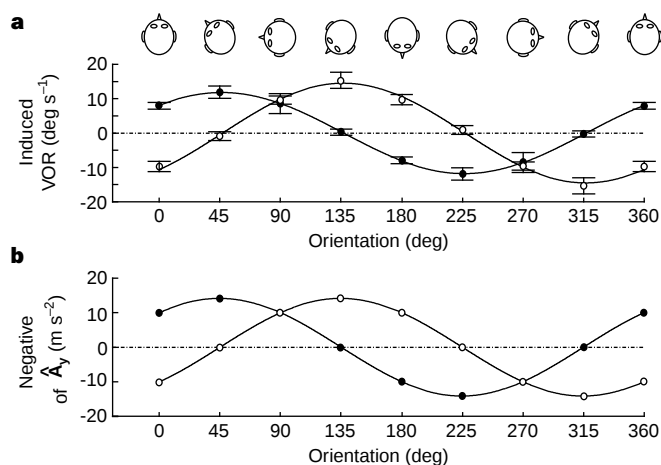


Figure 4 Comparison of induced VOR to theoretical estimate of linear acceleration. **a**, We plot the induced VOR component between 3 and 4 s after the completion of the tilt to allow time for the responses to stabilize. The mean value and standard error are plotted against subject orientation following both CW (○) and CCW (●) rotations. Least-square-error sinusoidal fits to the data are shown. **b**, The negative of the predicted magnitude of the inter-aural, linear acceleration component is plotted against subject orientation for 45° tilts in the estimated linear acceleration to the left (○) or right (●) of vertical. The negative is plotted to account for the 180° phase shift required of a compensatory response.

those used in this study showed that this paradigm induced an estimate of linear acceleration, evoking an induced VOR that contributed to the total VOR (though this contribution was not recognized at the time of publication). Another recent study²⁸ confirms these findings and proposes neural mechanisms that are identical to those included in the earlier modelling work.

Our results show that illusory tilt can be elicited by cues from the semicircular canals and that the VOR following post-rotational tilt depends systematically and predictably on both subject orientation and direction of rotation. Neither of these findings is consistent with theories suggesting that ambiguous gravito-inertial cues are resolved via simple high-pass or low-pass filters (peripheral or central). The findings presented here are consistent with our hypothesis that an internal model resolves the ambiguous gravito-inertial cues measured by graviceptors, with cues from the semicircular canals used to help interpret ambiguous gravito-inertial cues. Internal models may represent a general neural process to resolve sensory ambiguities^{2,4}, to synthesize information from disparate sensory modalities^{11,12,27}, and to combine efferent and afferent information^{6,7,10}.

Received 1 October 1998; accepted 22 February 1999.

- Young, L. in *Handbook of Physiology—The Nervous System* (ed. Darian-Smith, I.) 1023–1066 (American Physiological Society, Bethesda, Maryland, 1984).
- Ormsby, C. & Young, L. Integration of semicircular canal and otolith information for multisensory orientation stimuli. *Math. Biosci.* **34**, 1–21 (1977).
- Merfeld, D. M. & Young, L. R. The vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt. *Exp. Brain Res.* **106**, 111–122 (1995).
- Merfeld, D. M. Modeling the vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt. *Exp. Brain Res.* **106**, 123–134 (1995).
- Yasui, S. & Young, L. Perceived visual motion as effective stimulus to pursuit eye movement system. *Science* **190**, 906–908 (1975).
- Darlot, C., Zupan, L., Etard, O., Denise, P. & Maruani, A. Computation of inverse dynamics for the control of movements. *Biol. Cybern.* **75**, 173–186 (1996).
- Wolpert, D., Ghahramani, Z. & Jordan, M. An internal model for sensorimotor integration. *Science* **269**, 1880–1882 (1995).
- Paulin, M. G. The role of the cerebellum in motor control and perception. *Brain Behav. Evol.* **41**, 39–50 (1993).
- Bell, C., Bodznick, D., Montgomery, J. & Bastian, J. The generation and subtraction of sensory expectations within cerebellum-like structures. *Brain Behav. Evol.* **50** (suppl. 1), 17–31 (1997).
- Kuo, A. An optimal control model for analyzing human postural balance. *IEEE Trans. Biomed. Eng.* **42**, 87–101 (1995).
- Droulez, J. & Darlot, C. in *Attention and Performance XIII* (ed. Jeannerod, M.) 495–526 (Erlbaum, New York, 1989).
- Glasauer, S. Interaction of semicircular canals and otoliths in the processing structure of the subjective zenith. *Annl. N. Y. Acad. Sci.* **656**, 847–849 (1992).
- Rao, R. & Ballard, D. Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput.* **9**, 721–763 (1997).
- Lewald, J. & Ehrenstein, W. Influence of head-to-trunk position on sound lateralization. *Exp. Brain Res.* **121**, 230–238 (1998).
- Schwarz, C., Busetini, C. & Miles, F. Ocular responses to linear motion are inversely proportional to viewing distance. *Science* **245**, 1394–1396 (1989).
- Paige, G. & Tomko, D. Eye movement responses to linear head motion in the squirrel monkey. I. Basic characteristics. *J. Neurophysiol.* **65**, 1170–1182 (1991).
- Schwarz, C. & Miles, F. Ocular responses to translation and their dependence on viewing distance. I. Motion of the observer. *J. Neurophysiol.* **66**, 851–864 (1991).
- Wilson, V. & Melvill Jones, G. *Mammalian Vestibular Physiology* (Plenum, New York, 1979).
- Guedry, F. in *Handbook of Sensory Physiology* (ed. Kornhuber) 1–154 (Springer, New York, 1974).
- Stockwell, C. & Guedry, F. The effect of semicircular canal stimulation during tilting on the subsequent perception of the visual vertical. *Acta Otolaryngol.* **70**, 170–175 (1970).
- Dichgans, J., Held, R., Young, L. R. & Brandt, T. Moving visual scenes influence the apparent direction of gravity. *Science* **178**, 1217–1219 (1972).
- Benson, A. & Bodin, M. Comparison of the effect of the direction of the gravitational acceleration on post-rotational responses in yaw, pitch, and roll. *Aerospace Med.* **37**, 889–897 (1966).
- Goldberg, J. & Fernandez, C. Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. I. Resting discharge and response to constant angular accelerations. *J. Neurophysiol.* **34**, 635–660 (1971).
- Benson, A. in *Handbook of Sensory Physiology Vol. VI, Vestibular System Part 2: Psychophysics, Applied Aspects and General Interpretations*. (ed. Kornhuber, H.) 281–320 (Springer, Berlin, Heidelberg and New York, 1974).
- Fetter, M., Tweed, D., Hermann, W., Wohlend-Braun, B. & Koenig, E. The influence of head position and head reorientation on the axis of eye rotation and the vestibular time constant during postrotatory nystagmus. *Exp. Brain Res.* **91**, 121–128 (1992).
- Merfeld, D. M. Modeling human vestibular responses during eccentric rotation and off vertical axis rotation. *Acta Otolaryngol.* (suppl. 520), 354–359 (1995).
- Merfeld, D. M., Young, L., Oman, C. & Shelhamer, M. A multi-dimensional model of the effect of gravity on the spatial orientation of the monkey. *J. Vestibular Res.* **3**, 141–161 (1993).
- Angelaki, D., McHenry, M., Dickman, J. D., Newlands, S. & Hess, B. Computation of inertial motion: Neural strategies to resolve ambiguous otolith information. *J. Neurosci.* **19**, 316–327 (1999).

Acknowledgements. We thank S. Clark-Donovan and J. Roth for assistance throughout and T. Inglis, G. McCollum, S. Oster, P. Roberts, C. Wall and the NSI sensorimotor journal club for comments on the manuscript. This work was supported by grants from the NIH/NIDCD (D.M.M.), ESA (L.Z.), and NIH/NASA (R.J.P.).

Correspondence and requests for materials should be addressed to D.M.M. (e-mail: dan_merfeld@meei.harvard.edu).

A rhoptyr-protein-associated mechanism of clonal phenotypic variation in rodent malaria

P. R. Preiser*, W. Jarra*, T. Capiod*†‡ & G. Snounou§

Divisions of * Parasitology and † Neurophysiology and Neuropharmacology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

§ Department of Infection and Tropical Medicine, Imperial College School of Medicine, Northwick Park Hospital, Lister Unit, Harrow, Middlesex HA1 3UJ, UK

The recognition and invasion of host cells are mediated by components of the apical complex of the ookinete, sporozoite and merozoite stages of *Plasmodium* parasites¹. The paired rhoptries (organelles involved in host-cell recognition) in the apical complex contain many proteins of as-yet unknown function. In the rodent malaria agent *P. yoelii yoelii*, a multigene family codes for merozoite rhoptyr proteins of relative molecular mass 235,000 (p235 proteins)^{2,3}; these proteins are thought to determine the subset of erythrocytes that the parasites invade^{4,5}. Further support for this idea came from the identification of a region in p235 with weak but significant homology to reticulocyte-binding protein-2 of *P. vivax*^{6–8} and the demonstration that at least one p235 member binds to the erythrocyte surface membrane⁹. Here, using single, micromanipulated *P. yoelii* parasites, we describe a new mechanism of gene expression by which the merozoites originating from a single schizont each express a distinct member of this multigene family. We propose that this new type of clonal phenotypic variation provides the parasite with a survival strategy in the mammalian host; this strategy contributes to the observed chronicity of malarial infections. This phenomenon is genetically and functionally distinct from classical antigenic variation, which is mediated by the *var* multigene family of *P. falciparum*^{10–13}.

There are up to 50 copies of the *P. yoelii* p235 rhoptyr-protein genes (*Py235*) per genome, and there are at least 11 distinct *Py235* genes in this multigene family⁷. Furthermore, a region of the 3' ends of the *Py235* genes differs in the number or composition of a triamino-acid unit that it encodes⁷ (Fig. 1). Using an approach based on the polymerase chain reaction (PCR) with universal oligonucleotide primers spanning these variable sequences, we established that the *Py235* genes in the virulent YM and the avirulent, reticulocyte-restricted 17X lines of *P. y. yoelii* can be subdivided into at least four subgroups¹⁴. Transcripts of all these subgroups could be detected by reverse transcription with PCR (RT-PCR) in 17X parasites, whereas only three were transcribed in YM-parasite populations¹⁴. Restriction of YM parasites to reticulocytes might have been expected to result in a change of the transcription pattern of the *Py235* genes. However, no such changes could be observed by using RT-PCR to compare YM parasites in control infection with those restricted to reticulocytes by passive transfer of anti-p235 monoclonal antibodies¹⁴. In these experiments, the use of pooled messenger RNA from samples containing >1,000 parasites might have made it difficult to detect changes in the proportion of parasites expressing different members of the *Py235* gene family. Moreover, meaningful interpretation of such data depends on understanding the mechanism by which transcription of the different *Py235* gene members is regulated in individual parasites. To address this point we analyse here *Py235* transcription in micro-manipulated single erythrocytic parasites.

† Present address: INSERM U442, Université Paris-Sud, Bat. 443, Orsay 91405, France.

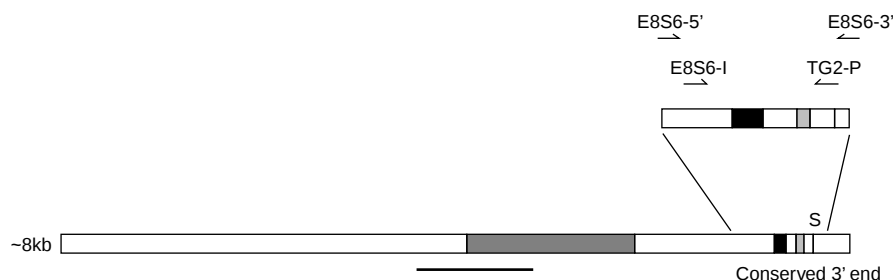


Figure 1 Structure of the *Py235* genes, highlighting areas of potential functional importance and the region analysed here. The cross-hatched area indicates the region of significant homology with the gene encoding *P. vivax* reticulocyte-binding protein-2 (ref. 7); the 1-kilobase region found in at least 11 different members of the *Py235* family is underlined³. Primers E8S6-3' and E8S6-5' span a region of ~590 base pairs, with the nested primer pair E8S6-I and TG2-P

covering ~400 base pairs located 5' of the stop codon (S). Both primer pairs cover the variable amino-acid (sequence DIN) repeat region (black) as well as the region encoding the transmembrane domain (speckled). The area of highest conservation is towards the 3' end of the gene; more variability between the members is observed in moving towards the 5'-end.

P. y. yoelii completes its erythrocytic asexual cycle in ~24 hours. As *P. y. yoelii* infections in mice are asynchronous, we used Percoll gradients to obtain fractions enriched for 17X parasites of morphologically distinct stages of maturity. We analysed three categories of parasite: first, the ring forms and early trophozoites; second, the middle to late trophozoites; and finally, the mature trophozoites and schizonts (see life cycle shown in Fig. 2). Parasites in the first two categories are characterized by a single nucleus, whereas in the third category, nuclear division having been initiated, parasites

contain 2–16 nuclei. Individual parasitized erythrocytes were isolated by single-cell micromanipulation and either DNA or RNA was extracted from each cell (see Methods). Three or four variant forms of the 3' end region were always observed when the DNA from a single parasite, irrespective of the fraction from which it was isolated, was analysed by PCR (Fig. 2, DNA panels). Using RT-PCR we then investigated the pattern of *Py235* transcription in single parasites at different stages of maturity. No transcripts could be detected in any of the parasites of the first category (that is, in

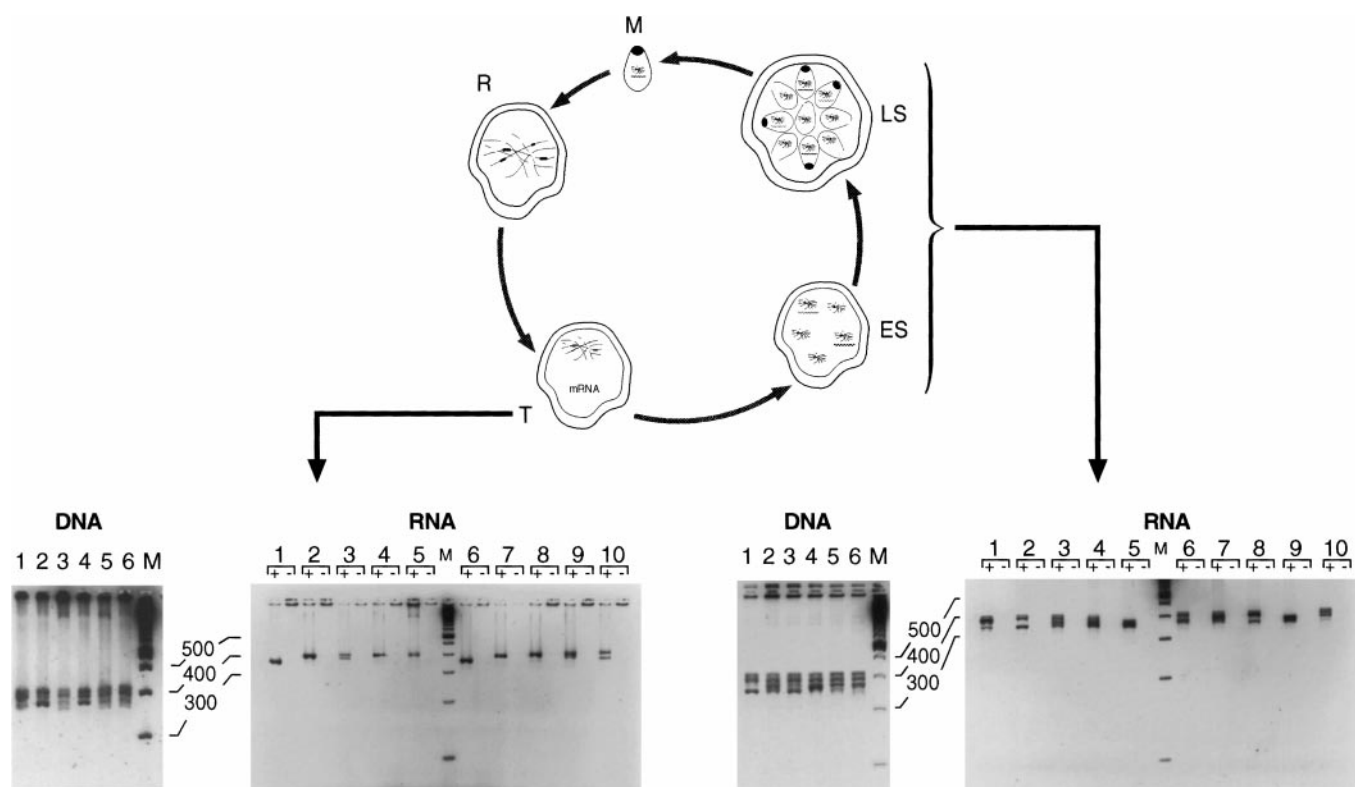


Figure 2 Analysis of PCR and RT-PCR products obtained from *P. y. yoelii* parasites at different morphological stages. Top, the asexual life cycle of *P. y. yoelii* parasites in erythrocytes. The different morphological stages are as follows: R, ring; T, trophozoite; ES, early schizont; LS, late schizont; M, merozoite. Chromosomes are represented in the parasites as strands carrying the different members of the *Py235* gene family. Our immature/mature-trophozoite fraction was predominantly composed of T parasites, with a minor component of ES parasites. The mature-parasite fraction was predominantly composed of multinucleated ES and LS parasites. Bottom, panels labelled 'DNA' show results of agarose-gel electrophoresis of nested PCR products, using as a template *P. y. yoelii* 17X genomic

DNA purified from six single, micromanipulated, T-fraction parasites (left) or ES/LS-fraction parasites (right) (see Methods). The lanes labelled 'M' show molecular-size markers (given in base pairs). The panels labelled 'RNA' show results of agarose-gel electrophoresis of the RT-PCR products obtained from the RNA of single micromanipulated parasites. The RNA was equally split and reverse transcriptase was included in one half (+) and omitted from the other half (-), thus allowing us to control for the possibility of DNA contamination. Results of the analysis for ten individual parasites from the T fraction (left) or from the ES/LS fraction (right) are shown.

freshly invaded erythrocytes), whereas ~50% of the parasites in the second category resulted in a positive RT-PCR result, and in most of these only a single transcript could be observed (Fig. 2, T-fraction RNA panels). The fact that the single transcripts found in the middle to late trophozoites differed from each other, and that the four variants were equally represented, indicates that there is no apparent predetermined order of gene transcription following invasion. The few instances in which two variants were found in the same parasite (Fig. 2, T-fraction RNA panel, samples 3, 9, 10) probably arose from a contamination of the fraction with older parasites (see Methods).

In contrast, >90% of the parasites from the third fraction showed multiple transcripts, often with the four variants present in a single parasite (Fig. 2, ES/LS fraction RNA panel). As the region analysed here is located at the 3' end of genes of about 8 kilobases in length, it is unlikely that the transcripts detected in the different parasite fractions are the result of low-level run-on transcription ('leaky transcription'). Expression of the *Py235* family is therefore initiated by the transcription of only one or a restricted subset of these genes as the parasite nears the first round of nuclear division. The increase in the number of genes becoming transcriptionally active appears to coincide with the replication of the parasite genome during schizogony, by the end of which multiple transcripts are present.

Individual schizonts, therefore, contain multiple transcripts of the *Py235* genes. We next determined whether this is also the case for single merozoites. We isolated individual mature schizonts by micromanipulation, and the merozoites within these schizonts were then de-aggregated by a brief sonication. Aliquots containing single merozoites were obtained by limiting dilution of the resulting material (see Methods). RT-PCR analysis showed that only a single transcript could be detected in individual merozoites (Fig. 3). Moreover, merozoites derived from a single schizont were found to contain different transcripts (Fig. 3, samples 5–9). We confirmed this result by sequencing all the RT-PCR products (data not shown). To exclude the possibility that these results are due to PCR artefacts, we performed control experiments in which two or more of the genes bearing the different 3'-end variants were mixed in equal proportions. We could detect all the variants even at very low template concentrations (1–100 copies). In mixtures containing a total of 100 or 1,000 templates, in which one variant predominates, the minor variant was only occasionally detected at a 1:100 ratio, frequently detected at 1:25, and always at 1:10 (data not shown). We therefore conclude that each merozoite predominantly contains transcripts of a single gene.

Transcription from all the members of the *Py235* family in a single parasite explains our failure to detect changes in the transcription pattern following antibody selection of reticulocyte-invading *P. y. yoelii* YM parasites¹⁴. Our observations indicate that the reduced virulence of YM parasites, observed following passive transfer of antibodies or immunization with purified p235 proteins⁴, is due not to a transcriptional switch of the *Py235* genes, but to immune evasion by a subset of merozoites. Immunological reagents and functional assays specific to different members of the p235 family will be required to provide formal confirmation of our conclusions.

Our results contrast with observations on the transcription of the *var* gene family of *P. falciparum*. This family codes for the antigenically variable PfEMP1 surface proteins, which mediate parasite cytoadherence, a phenomenon that underlies the pathology of the infection^{11–13}. Minor quantities of multiple *var* gene transcripts are detected in ring-stage parasites, and as these organisms mature to form schizonts only one transcript predominates¹⁵. It therefore appears that transcription of these two multigene families is controlled by quite distinct mechanisms.

Our data suggest a model of transcription (Fig. 4) in which the number of *Py235* genes transcribed correlates directly with the number of nuclei present in the parasite. Transcription of one member of the gene family is initiated at the late trophozoite stage. As the parasite matures to form a multinucleated schizont,

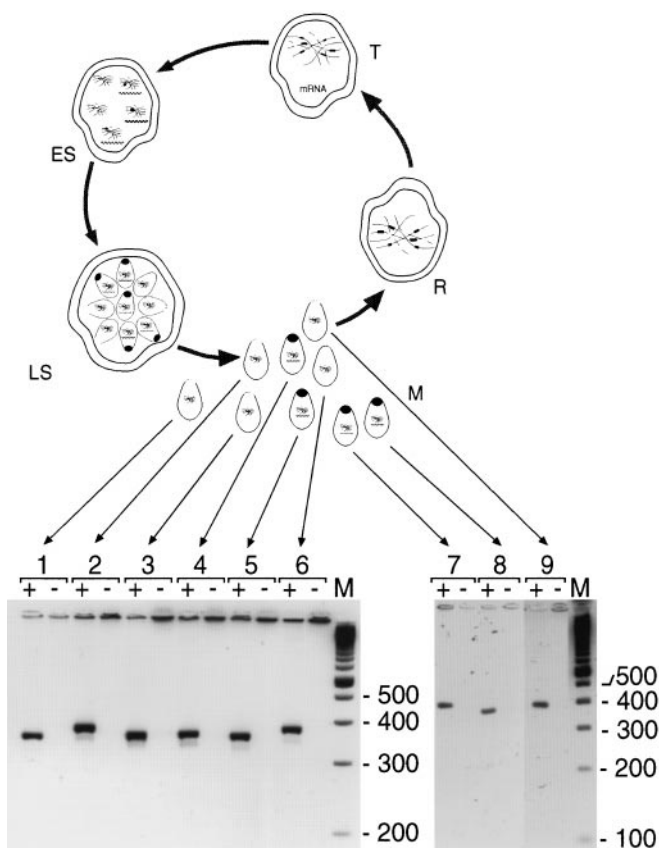


Figure 3 Representative results from RT-PCR analysis of RNA purified from single merozoites. RNA from single merozoites was split into equal halves and analysed by RT-PCR in the presence (+) and absence (–) of reverse transcriptase. Merozoites 3 + 4, 5 + 6, and 7–9 were each derived from a single schizont. Sequencing of cloned products showed that the transcripts detected in merozoites 1, 5 and 9 were identical to each other; the transcripts in merozoites 3, 4 and 7 were also identical, as were those in merozoites 6 and 8. The lane labelled 'M' shows a 100-base-pair marker ladder.

further *Py235* genes become transcriptionally active, presumably by a stochastic process. It is, however, difficult to determine through micromanipulation whether the particular gene first transcribed in the trophozoite corresponds to the one transcriptionally active in the merozoite from which it originated. It will be interesting to determine whether the mRNA produced by each of the genomes present in the syncytial schizont remains associated with its respective nucleus before segregation to individual merozoites.

The efficiency with which *Plasmodium* parasites invade erythrocytes can be considered as an important virulence factor. Thus, in the natural host, infections by parasites restricted to reticulocytes (such as *P. vivax* and *P. ovale* in man, and *P. berghei* in rodents) rarely reach high levels and are consequently not associated with substantial mortality. The fact that multiple pathways for the invasion of erythrocytes are available is advantageous to the parasite, because this allows it to maintain an adequate multiplication rate despite the acquisition of immunity against a given receptor and the heterogeneity of the erythrocytes circulating in the mammalian host.

The pattern of transcription that we describe here indicates a mechanism by which the parasite can adapt efficiently to variations in the host environment. If most parasites express only one type of erythrocyte receptor, parasite populations will be vulnerable to erythropoietic changes or to immune pressure in the host. Similarly, if each merozoite were to express multiple p235 proteins, a higher proportion of these merozoites would be inactivated by specific

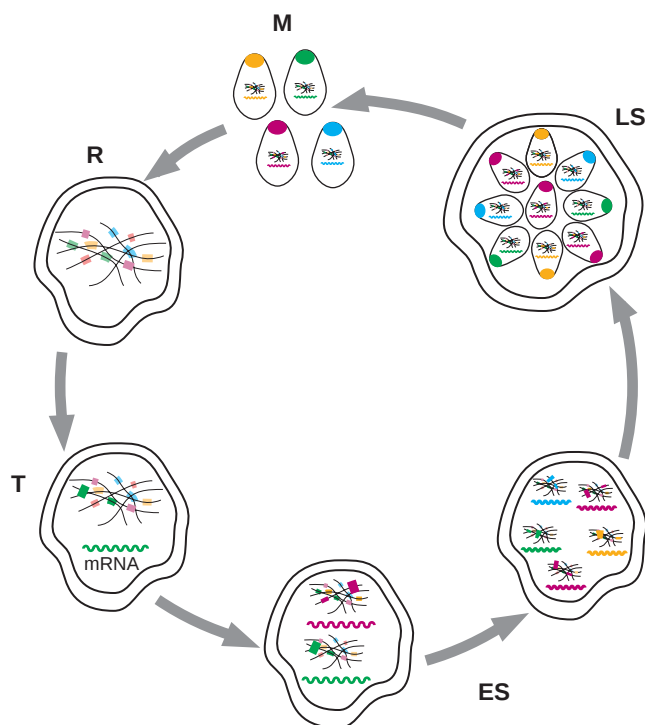


Figure 4 Proposed model for *Py235* transcription in *P. y. yoelii*, based on analysis of the four differently sized alleles used here (these have each been assigned a different colour). No transcription could be detected at the ring stage. At the trophozoite stage, only one transcript is seen; as the first nuclear division occurs, the second nucleus can transcribe a different message, leading to the presence of two different transcripts in the very early schizonts. As schizonts mature, 8–16 nuclei are formed, each capable of transcribing a different member of the *Py235* gene family, leading to the detection of all four transcript types in a single late-stage schizont. Each merozoite contains only one transcript and probably, therefore, only one type of p235 protein in its rhoptries.

immune responses. Thus, by ensuring that a single schizont produces merozoites each with a distinct specificity, the ability of the parasite to maintain the infection, and thus to increase its potential for transmission, is optimized. An important characteristic of this type of phenotypic variation is the subtle biological flexibility and adaptability it confers to an organism. The potential diversity of multigene families is thus made available to the progeny of each cell at each cycle of multiplication. Homologues of the p235 proteins have been found in other malaria species, including *P. vivax* of man^{6,7,16}, and are also likely to be found in *P. falciparum*. It will be interesting to determine whether a similar mechanism of host-cell selection occurs in the other medically and economically important parasitic apicomplexan species. □

Methods

Parasite preparation and micromanipulation. Female BALB/c mice, bred specific pathogen-free at the National Institute for Medical Research, were housed as described¹⁷. Parasites were obtained from the cloned *P. yoelii* 17X line and maintained either as cryopreserved stabiles¹⁷ or by weekly syringe passage in mice. Experimental animals were infected by intraperitoneal injection of parasitized erythrocytes; the parasitaemia was monitored on thin blood smears and infected blood was collected from these animals as described¹⁷.

Very mature trophozoites and schizonts were separated from immature parasites on a 60% Percoll cushion. Immature trophozoites/mature trophozoites were obtained from a discontinuous gradient composed of equal volumes of 60%, 70% and 80% Percoll. The immature fraction (≤ 2 nuclei per parasite) consisted of about 27% immature trophozoites, 62% mature trophozoites, and 11% early schizonts. No late-stage schizonts (> 2 nuclei

per schizont) were detected even after an extensive microscopic survey. The mature-parasite fraction contained 2% immature trophozoites, 27% mature trophozoites and 71% schizonts. Single parasites were selected under phase-contrast optics and micromanipulated using micropipettes pulled from GC150-F borosilicate capillary tubes (Clark Electromedical Instruments).

Isolation of single merozoites. Mature schizonts were isolated individually by micromanipulation from an enriched Percoll gradient fraction. Each schizont was ruptured by brief sonication to release intact merozoites and separated into 20 aliquots. The precise release conditions for the production of intact merozoites were predetermined using schizonts prestained with 4,6-diamidino-2-phenylindole and viewed by ultraviolet microscopy. The number of merozoites per mature *P. yoelii* schizont ranges between 8 and 16, so PCR analysis of 10 of the 20 samples would be expected to give 4–8 positive results, and this was routinely seen. However, complete lysis of all the merozoites by excessive sonication gave positive results in every sample (data not shown). The remaining merozoites from preparations that gave between five and eight (out of ten) positive results were used for RNA extraction and RT-PCR. Only about 25% of all the merozoites analysed by RT-PCR gave a positive result, probably because of the low amount of RNA remaining in them.

RNA and DNA preparation from single parasites. Large-scale RNA and DNA extractions were carried out as described¹⁴. RNA from single parasites was isolated using the protocol supplied with the Clontech microscale total RNA separator kit. Before using the reverse-transcriptase step, the RNA was treated with DNase and reprecipitated.

DNA from single parasites was extracted using a similar protocol. The parasites were lysed in 90 μ l denaturing buffer; 10 μ l poly (I) (2 mg ml⁻¹) and 10 μ l 2 M sodium acetate (pH 4.2) were then added and the sample was extracted with Tris-HCl (pH 8.0) equilibrated with phenol:chloroform. DNA was then precipitated with isopropanol.

PCR and RT-PCR. PCR and RT-PCR of material obtained from a large-scale DNA or RNA preparation were carried out as described¹⁴. For the PCR of DNA from a single parasite, the DNA pellet obtained was directly resuspended in PCR buffer containing primers E8S6-3' and E8S6-5' and thermal cycling was carried out as described¹⁴. For the nested PCR reaction, 2 μ l of the first reaction were transferred directly to a second PCR reaction containing primers E856-I and TG2-P (ref. 14).

RNA from a single parasite was resuspended in 23 μ l H₂O and 2 μ l primer E8S6-3' (100 ng μ l⁻¹) was added. Hybridization and reverse transcription were carried out according to the manufacturer's protocol (Superscript II RNaseH reverse transcriptase, Gibco). The sample was split in two; 200 units of reverse transcriptase were added to one half, and enzyme was omitted from the other half. The samples were incubated at 42 °C for 30 min, followed by 3 min at 95 °C to inactivate the enzyme. To each tube, 30 μ l of a PCR buffer (containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 3 mM MgCl₂, 0.001% (w/v) gelatin, 1 mM of each dNTP, 100 ng E8S6-3', 200 ng E8S6-5', and 2 units of AmpliTaq DNA polymerase (Perkin Elmer)) was added. Amplification was carried out as described¹⁴. A 2- μ l aliquot from the first reaction was directly transferred to the nested PCR reaction containing primers E8S6-I and TG2-P¹⁴. Samples from all of the PCR and RT-PCR reaction were analysed on 3% Mota Phore® agarose gels (FMC Bioproducts).

Received 5 October 1998; accepted 5 February 1999.

- Sam-Yellow, T. Y. Rhoptry organelles of the Apicomplexa: their role in host cell invasion and intracellular survival. *Parasitol. Today* **12**, 308–316 (1996).
- Keen, J. et al. Identification of the gene for a *Plasmodium yoelii* rhoptry protein. Multiple copies in the parasite genome. *Mol. Biochem. Parasitol.* **42**, 241–246 (1990).
- Borre, M. B. et al. Multiple genes code for high-molecular-mass rhoptry proteins of *Plasmodium yoelii*. *Mol. Biochem. Parasitol.* **70**, 149–155 (1995).
- Holder, A. A. & Freeman, R. R. Immunization against blood-stage rodent malaria using purified parasite antigens. *Nature* **294**, 361–364 (1981).
- Freeman, R. R., Trejdosiewicz, A. J. & Cross, G. A. M. Protective monoclonal antibodies recognising stage-specific merozoite antigens of a rodent malaria parasite. *Nature* **284**, 366–368 (1980).
- Galinski, M. R. et al. A reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. *Cell* **69**, 1213–1226 (1992).
- Keen, J. K. et al. A gene coding for a high-molecular mass rhoptry protein of *Plasmodium yoelii*. *Biochem. Parasitol.* **65**, 171–177 (1994).
- Sinha, K. A. et al. Comparison of two members of a multigene family coding for high-molecular mass rhoptry proteins of *Plasmodium yoelii*. *Mol. Biochem. Parasitol.* **76**, 329–332 (1996).
- Ogun, S. A. & Holder, A. A. A high molecular mass *Plasmodium yoelii* rhoptry protein binds to erythrocytes. *Mol. Biochem. Parasitol.* **76**, 321–324 (1996).
- Brown, K. N. & Brown, I. N. Immunity in malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. *Nature* **208**, 1286–1288 (1965).
- Smith, J. D. et al. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in

- antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110 (1995).
12. Su, X. Z. *et al.* The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89–100 (1995).
 13. Baruch, D. I. *et al.* Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82**, 77–87 (1995).
 14. Preiser, P. R. & Jarra, W. *Plasmodium yoelii*: differences in the transcription of the 235 kDa rhoptry protein multi-gene family in lethal and non lethal lines. *Exp. Parasitol.* **89**, 50–57 (1998).
 15. Chen, Q. *et al.* Developmental selection of var gene expression in *Plasmodium falciparum*. *Nature* **394**, 392–395 (1998).
 16. Galinski, M. R. & Barnwell, J. W. *Plasmodium vivax*: merozoites, invasion of reticulocytes and considerations for malaria vaccine development. *Parasitol. Today* **12**, 20–29 (1996).
 17. Jarra, W. & Brown, K. N. Protective immunity to malaria: studies with cloned lines of rodent malaria in CBA/CA mice. IV. The specificity of mechanisms resulting in crisis and resolution of the primary acute phase parasitaemia of *Plasmodium chabaudi chabaudi* and *P. yoelii yoelii*. *Parasite Immunol.* **11**, 1–13 (1989).

Acknowledgements. We thank C. Magnus and G. H. Mitchell for help with micromanipulation, and K. N. Brown, A. A. Holder, S. Khan, M. Knight, R. J. Wilkinson and R. J. M. Wilson for helpful advice and critical reading of the manuscript.

Correspondence and requests for material should be addressed to P.R.P. (e-mail: ppreise@nimr.mrc.ac.uk).

Homeobox gene *Nkx2.2* and specification of neuronal identity by graded Sonic hedgehog signalling

J. Briscoe*, L. Sussel†, P. Serup‡, D. Hartigan-O'Connor†, T. M. Jessell, J. L. R. Rubenstein† & J. Ericson*

Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA

†Nina Ireland Laboratory of Developmental Neurobiology,

University of California at San Francisco, 402 Parnassus Avenue, LPP1 Box 0984, San Francisco, California 94143-0984, USA

* These authors contributed equally to this work.

During vertebrate development, the specification of distinct cell types is thought to be controlled by inductive signals acting at different concentration thresholds¹. The degree of receptor activation in response to these signals is a known determinant of cell fate², but the later steps at which graded signals are converted into all-or-none distinctions in cell identity remain poorly resolved. In the ventral neural tube, motor neuron and interneuron generation depends on the graded activity of the signalling protein Sonic hedgehog (Shh)^{3–5}. These neuronal subtypes derive from distinct progenitor cell populations that express the homeodomain proteins *Nkx2.2* or *Pax6* in response to graded Shh signalling^{6,7}. In mice lacking *Pax6*, progenitor cells generate neurons characteristic of exposure to greater Shh activity^{6,8}. However, *Nkx2.2* expression expands dorsally in *Pax6* mutants⁶, raising the possibility that *Pax6* controls neuronal pattern indirectly. Here we provide evidence that *Nkx2.2* has a primary role in ventral neuronal patterning. In *Nkx2.2* mutants, *Pax6* expression is unchanged but cells undergo a ventral-to-dorsal transformation in fate and generate motor neurons rather than interneurons. Thus, *Nkx2.2* has an essential role in interpreting graded Shh signals and selecting neuronal identity.

We analysed the pattern of neurogenesis in the ventral spinal cord and hindbrain of mice containing a mutation in the *Nkx2.2* gene⁹, focusing on two ventral progenitor cell populations. The ventral-most neuronal progenitor domain, defined as the region dorsolateral to HNF3 β floorplate cells¹⁰ and ventral to *Pax6*^{on} progenitors, is initially characterized by expression of *Nkx2.2* and the related *Nkx2.9* gene^{11–13} (Fig. 1a–e). This domain corresponds to the 'x' region¹⁴. From e10.5, high *Nkx2.2* expression persists but in the spinal cord the expression of *Nkx2.9* by 'x'-domain progenitors is

almost extinguished¹³ (Fig. 1b, c, e, f). More dorsal progenitor cells express low levels of *Pax6* but not of *Nkx2.2* or *Nkx2.9* (Fig. 2a, b). Both progenitor cell populations express *Nkx6.1* (Fig. 2c, h)¹⁵.

To test whether *Nkx2.2* helps to establish or maintain the 'x' progenitor domain, we monitored the expression of *Nkx2.9*, *Nkx6.1*, HNF3 β and *Pax6* in the spinal cord of mice lacking *Nkx2.2*. The pattern of expression of *Nkx2.9* and *Nkx6.1* over the period e9.5 to e12.5 was similar in wild-type, heterozygote and *Nkx2.2* mutant mice (Figs 1d, e, g, h and 2c, h; also data not shown). In addition, the ventral boundary of the *Pax6* progenitor cell domain was unchanged in *Nkx2.2* mutants (Fig. 2a, b, f, g), as were the domains and amount of expression of HNF3 β and Shh (Figs 2b, g and 4a, g; also data not shown). Moreover, the number of progenitor cells within the 'x' domain was not altered in *Nkx2.2* mutants (Fig. 2b, c, g, h, and data not shown). Thus, *Nkx2.2* function is not required to establish or maintain the 'x' progenitor domain, although *Nkx2.2* and *Nkx2.9* may have redundant activities. As Shh signalling is required for the establishment of the 'x' domain⁶, we investigated whether Shh might also maintain the 'x' domain, even in the absence of *Nkx2.2* function, by culturing ventral neural tube/floorplate (vf) explants derived from e10.5 *Nkx2.2* mutant or wild-type embryos in the presence of an anti-Shh antibody¹⁶. The 'x' progenitor domain was maintained in wild-type (Fig. 2l, m, q, r) and *Nkx2.2* mutant [vf] explants (Fig. 2n, o, s, t) in the presence of anti-Shh IgG (Fig. 2k, p). Thus, the 'x' progenitor domain, once established, can be maintained without Shh signalling or *Nkx2.2* activity.

We next tested whether neurons that derive from 'x'-domain progenitors are affected by the loss of *Nkx2.2*. In the spinal cord,

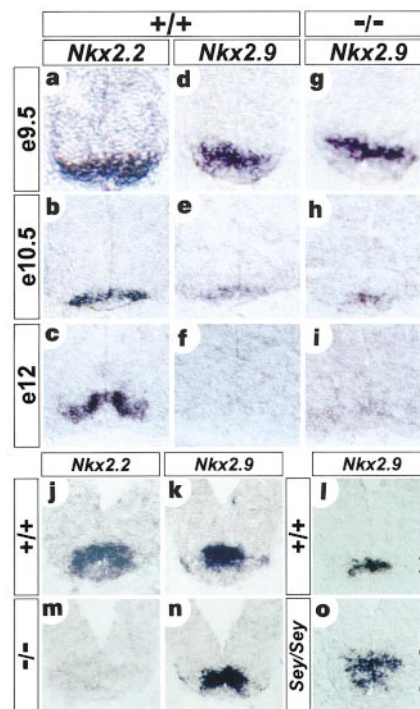


Figure 1 *Nkx2.2* and *Nkx2.9* are expressed in the 'x' domain of the neural tube. **a–c**, *Nkx2.2* expression in the 'x' domain of wild-type e9.5 (**a**), e10.5 (**b**) and e12 (**c**) mouse spinal cord. **d–f**, *Nkx2.9* is expressed in the 'x' domain of the spinal cord of wild-type (**d–f**) and *Nkx2.2* mutant (**g–i**) embryos at e9.5 (**d, g**) and downregulated by e10.5 (**e, h**). No *Nkx2.9* expression is detected at e12 (**f, i**). **j, k**, In the hindbrain *Nkx2.2* (**j**) and *Nkx2.9* (**k**) are expressed in the 'x' domain of wild-type e10.5 embryos. **m**, *Nkx2.2* expression is not detected in *Nkx2.2* mutant embryos, whereas in the hindbrain *Nkx2.9* (**n**) expression persists. **l, o**, The domain of *Nkx2.9* expression in the hindbrain of *Sey/Sey* embryos (**o**) expands dorsally with respect to the wild-type embryos (**l**).

‡ Present address: Hagedorn Research Institute, Niels Steensens Vej 6, DK-2820, Gentofte, Denmark.

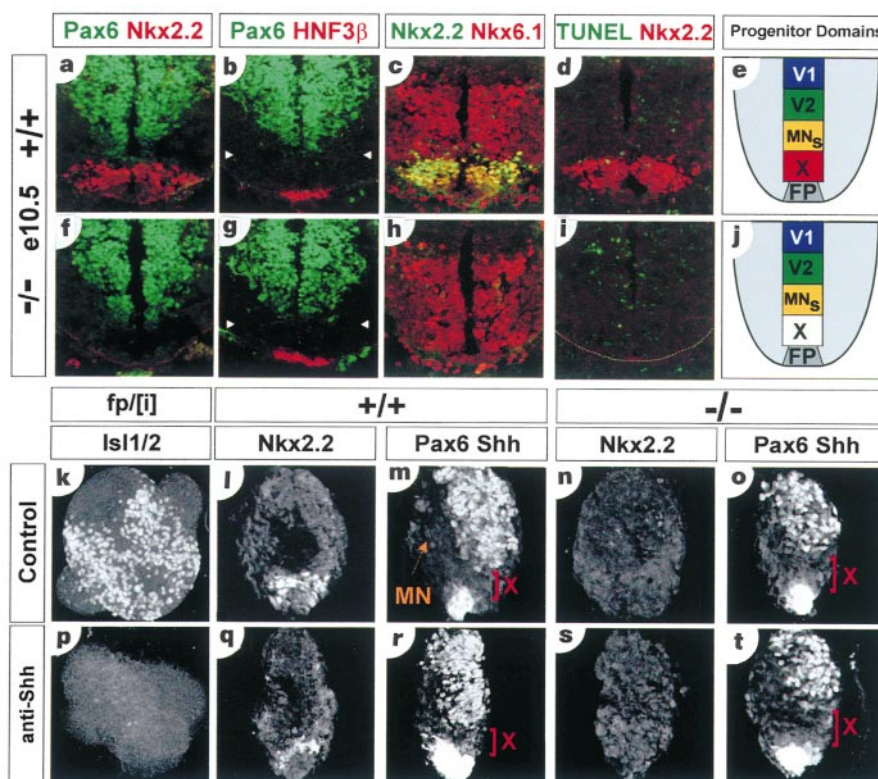


Figure 2 *Nkx2.2* activity is not required for the establishment of progenitor cell domains in the ventral neural tube. **a–j**, The 'x' domain is present in embryos lacking *Nkx2.2* activity. Expression of transcription factors in e10.5 wild-type (**a–c**) and *Nkx2.2* mutant (**f–h**) embryos. Pax6 but not *Nkx2.2* is expressed by mutant embryos (compare **a** and **f**). The lack of *Nkx2.2* activity does not result in a ventral expansion in the expression of Pax6 (**b, g**) (arrowheads indicate ventral boundary of Pax6 expression). *Nkx6.1* expression is similar in wild-type (**c**) and *Nkx2.2* mutant (**h**) embryos. The expression of Shh is similar in *Nkx2.2* mutant and wild-type embryos (Fig. 4a, g). The ventral limit of Pax7 is unchanged in *Nkx2.2* mutant embryos (data not shown), as is the number of V1 and V2 neurons (Fig. 3, and data not shown). **d, i**, The number of TUNEL-labelled cells in the 'x' domain of *Nkx2.2* mutant embryos (**i**) is similar to wild-type controls (**d**). **e, j**, Summary diagram showing progenitor domains in the ventral neural tube of wild-type and *Nkx2.2*

mutant embryos. **k–t**, The 'x' domain is maintained in the absence of Shh signalling or *Nkx2.2* expression. **k, p**, Floorplate from an e10.5 mouse embryo induces *Isl1/2*^{on} motor neurons in conjugates with [i] regions of stage-10 chick neural tube (**k**). This activity is blocked by anti-Shh IgG (25 µg ml⁻¹) (**p**). Anti-Shh IgG also blocks the induction of *Nkx2.2* in neural plate explants from e8.5 mouse embryos (not shown). Ventral neural tube/floorplate explants from 33–36 somite wild-type (**l, m, q, r**) or *Nkx2.2* mutant (**n, o, s, t**) embryos were cultured in the presence (**l–o**) or absence (**q–t**) of anti-Shh IgG for 24 h. Explants were labelled for expression of *Nkx2.2* (**l, q, n, s**) or Pax6 and Shh (**m, r, o, t**). All explants are orientated with the floorplate ventral. Motor neurons do not express Pax6 (orange arrow in **m**). The 'x' domain, defined by the region between Shh^{on} floor plate cells and Pax6^{on} progenitors (red brackets), is evident in both wild-type and *Nkx2.2* mutant embryos, irrespective of the presence of anti-Shh IgG.

'x'-domain progenitors give rise to a class of ventral neurons that we term V3 neurons (Fig. 3a–d) and which express the basic helix–loop–helix (bHLH) genes *Sim1* and *Ngn3* (refs 17, 18) (Fig. 3b, e). The expression of *Ngn3* and *Sim1* is first detected at ~e10.5 (Fig. 3e, f), a time when the expression of *Nkx2.9* has been markedly downregulated by 'x'-domain progenitors in the spinal cord (Fig. 1e). To determine whether *Nkx2.2* is required for the generation of V3 neurons, we analysed the expression of *Ngn3* and *Sim1* in *Nkx2.2* mutants between e10.5 and e12.5. At e10.5, the number of *Ngn3*^{on}, *Sim1*^{on} cells in *Nkx2.2*-mutant embryos was markedly reduced (Fig. 3e, f, i, j). By e12.5, there was a virtual absence of *Ngn3*^{on}, *Sim1*^{on} neurons (Fig. 3g, k; data not shown). The transient expression of *Nkx2.9* by 'x'-domain progenitors in *Nkx2.2* mutants (Fig. 1d–i) may account for presence, initially, of a small number of V3 neurons (see below). We detected no increase in the number of TUNEL⁺ cells in the 'x' domain (Fig. 2d, i), indicating that the loss of V3 neurons is not due to apoptosis. Thus, *Nkx2.2* activity is required for the generation and possibly also maintenance of V3 neurons (Fig. 3h, l).

In wild-type mice, the generation of somatic motor neurons is excluded from *Nkx2.2*^{on} 'x'-domain progenitors⁶. The dorsal expansion of *Nkx2.2* expression in Pax6 mutants is accompanied by the repression of somatic motor neuron differentiation⁶, suggesting that the expression of *Nkx2.2* within 'x'-domain progenitors sup-

presses their potential for somatic motor neuron generation. We therefore examined whether somatic motor neurons are generated from 'x'-domain progenitors in *Nkx2.2* mutants. The generation of somatic motor neurons was defined by co-expression of the homeodomain proteins *Isl1*, *Isl2*, *Lim3* and *HB9* (refs 6, 19–21). At e10.5, there was a 21% increase in the number of somatic motor neurons (Fig. 4a, g; see legend for quantification) and many motor neurons were located next to the floorplate, within the 'x' domain (Fig. 4a–m). The dorsal boundary of somatic motor neuron generation (Fig. 4a–c, g–i) and the number of V2 neurons (Fig. 4d, j; data not shown) were unchanged.

The detection of somatic motor neurons adjacent to the floorplate could reflect either their generation from 'x'-domain progenitors or their ventral migration from a more dorsal position of origin. To resolve this, we analysed the expression of *Lim3*, a LIM homeodomain protein that defines motor neuron progenitors as well as post-mitotic motor neurons^{6,19–21}. In *Nkx2.2* mutants examined at e10.5–e12.5, *Lim3*^{on} and *Isl1/2*^{off} motor neuron progenitors were located medially, within the ventricular zone of the 'x' domain (Fig. 4j, o) but were excluded from this domain in wild-type embryos (Fig. 4d, n). In *Nkx2.2* mutants, there was also an increase in the number of somatic motor neurons located at the border of the ventricular zone within the 'x' domain (Fig. 4n–p), a position that appears to reflect their migration laterally from the ventricular

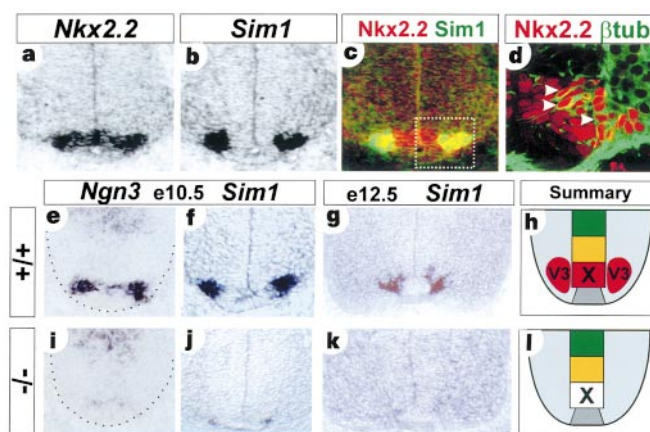


Figure 3 *Nkx2.2* activity is required for the generation of V3 neurons. **a–c**, *Sim1^{on}* neurons are generated from 'x' domain progenitors. Expression of *Nkx2.2* (**a**) and *Sim1* (**b**) in adjacent sections of e10.5 spinal cord. **c**, Superimposition of images **a** and **b** shows that 'x'-domain cells lateral to the ventricular zone co-express *Nkx2.2* and *Sim1*. **d**, Lateral *Nkx2.2^{on}* cells also express the neuronal specific β -tubulin isoform (arrowheads). Box in **c** shows the approximate position of the image in **d**. These cells are defined as neurons by expression of the neuronal specific β -tubulin isoform (**d**) and as V3 neurons by the selective expression of the genes encoding the bHLH transcription factor *Ngn3* (**e**); and the PAS bHLH

protein *Sim1* (**f**). **e–l**, Expression at e10.5 of *Ngn3* (**e**, **i**) and *Sim1* (**f**, **j**) in the spinal cord of wild-type (**e**, **f**) and *Nkx2.2* mutant (**i**, **j**) embryos. The expression of *Ngn3* and *Sim1* by cells in the 'x' domain is markedly reduced in *Nkx2.2* mutant embryos. However, dorsal expression of *Ngn3* is unaffected. By e12.5 (**g**, **k**), *Sim1* expression is lost in *Nkx2.2* mutant embryos (compare **g**, **k**). The progressive decrease in V3 neuron number between e11 and e12.5 argues that *Nkx2.2* may also be required for the maintenance of the V3 neuron phenotype. **h**, **l**, Summary diagrams showing the requirement for *Nkx2.2* in the generation of V3 neurons.

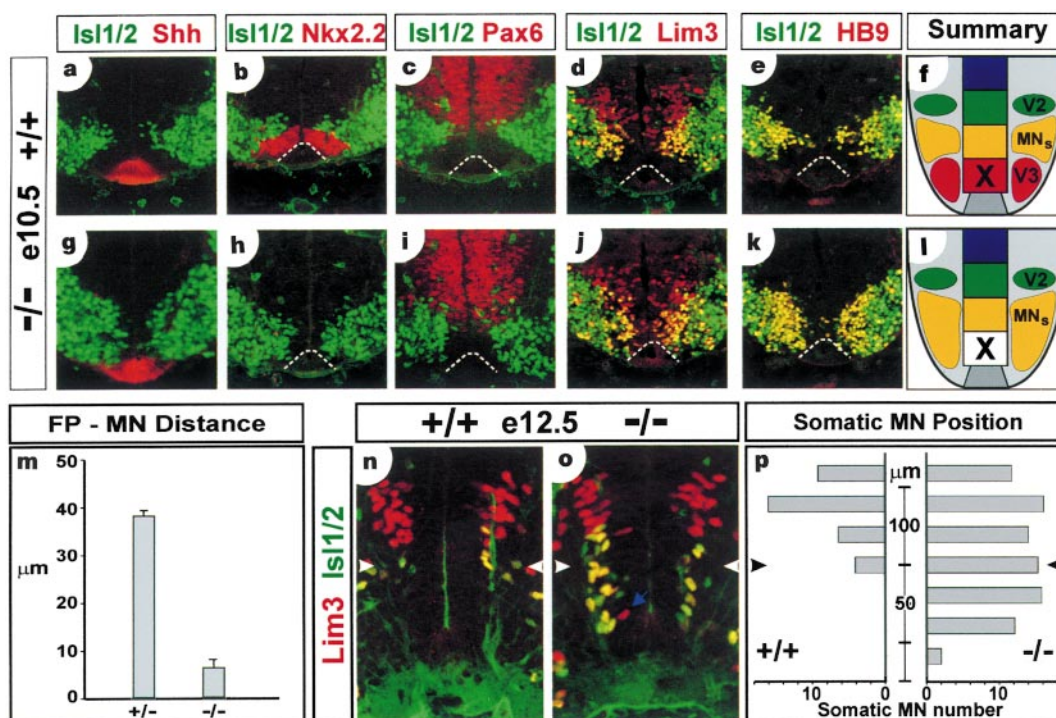


Figure 4 *Nkx2.2* represses somatic motor neuron generation in the 'x' domain of the spinal cord. **a–f**, Locations of somatic motor neurons at forelimb levels of wild-type (**a–f**) and *Nkx2.2* mutant (**g–l**) embryos at e10.5. Dotted line indicates the extent of the floorplate. *Nkx2.2^{on}* progenitors do not generate *Isl1/2^{on}* motor neurons (**b**, **h**). The gap between the floorplate and somatic motor neurons in wild-type mice (**a**) is missing in *Nkx2.2* mutant embryos (**g**). The ventral expression in somatic motor neurons results in the generation of motor neurons from *Pax6^{off}* progenitors (**c**, **i**). Ectopic motor neurons express *Lim3* (**d**, **j**), *HB9* (**e**, **k**) and *Isl1/2* (**c**, **i**). Number of somatic MNs: *Nkx2.2*^{+/–}: 160 ± 4; *Nkx2.2*^{–/–}: 202 ± 6 neurons per ventral quadrant *n* = 6 sections; mean ± s.e.m. Number of V2 neurons: *Nkx2.2*^{+/–}: 45 ± 9; *Nkx2.2*^{–/–}: 46 ± 13; neurons per ventral quadrant; *n* = 6 sections. **f**, **l**, Summary diagram showing the ventral expansion in somatic motor neuron generation in *Nkx2.2* mutant embryos. In *Nkx2.2* mutants examined

at e9.5, there was no change in the number or position of generation of somatic motor neurons (data not shown). This may reflect the overlap in expression of *Nkx2.9* and *Nkx2.2* (Fig. 1). Alternatively, the small dorsoventral extent of the 'x' domain at this early stage may obscure the ventral expansion of the domain of somatic motor neuron generation. **m**, Mean distance (μ m) between *Shh^{on}* floorplate cells and the most ventral *Isl1/2^{on}* motor neurons in wild-type and *Nkx2.2* mutant embryos at e10.5. At e12.5, *Lim3^{on}/Isl1/2^{on}* somatic motor neurons are present adjacent to the ventricular zone in wild-type embryos (**n**). There is a marked ventral expansion in the production of *Lim3^{on}*, *Isl1/2^{on}* motor neurons in *Nkx2.2* mutant embryos (**o**). Arrowheads mark the ventral limit of *Pax6* expression; the blue arrow indicates a *Lim3^{on}* motor neuron progenitor in the 'x' domain. **p**, Newly generated motor neurons as a function of distance from the ventral midline in e12.5 embryos.

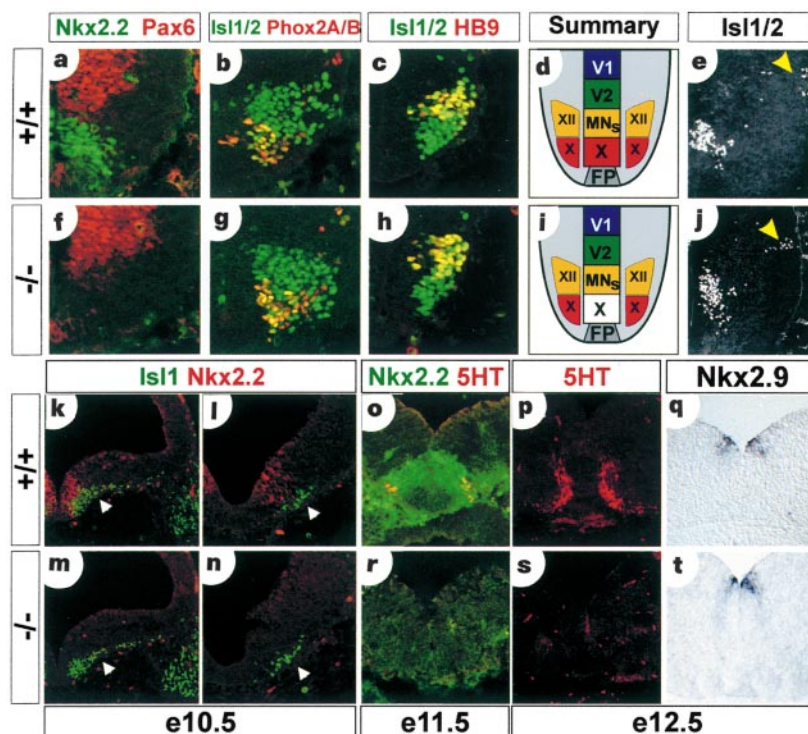


Figure 5 *Nkx2.2* is required for the generation of hindbrain serotonergic neurons but not visceral motor neurons. **a–j**, Position of generation of motor neurons at the r8 level of the hindbrain. In wild-type e10.5 embryos (**a–c**) *Nkx2.2*^{on} progenitors (**a**) generate *Isl1/Phox2A/B*^{on} visceral motor neurons (**b**), whereas *Pax6*^{ow} progenitors generate *Isl1/2*^{on} HB9^{on} somatic motor neurons (**c**). In *Nkx2.2* mutant embryos, 'x'-domain progenitors do not express *Nkx2.2* (**f**) but visceral (**g**) and somatic (**h**) motor neurons are generated in normal numbers and position. **d, i**, Summary diagram of visceral and somatic motor neuron generation in wild-type and *Nkx2.2* mutant embryos. **e, j**, Vagal and hypoglossal motor neurons migrate to their correct positions (yellow arrowheads) in wild-type (**e**) and *Nkx2.2* mutant (**j**) embryos. Additionally, whole-mount labelling of e11 wild-type and *Nkx2.2* mutant embryos using anti-neurofilament 2H3 antibody⁶ revealed that the hypoglossal motor nerve (XII) and vagus (X) and spinal accessory (XI) nerves are present (data not shown). The position of origin, transcription factor profile, and

migratory pattern of facial/glossopharyngeal, abducens, trochlear and oculomotor neurons is unchanged in *Nkx2.2* mutant embryos (data not shown). **k–n**, Trigeminal and trochlear *Isl1*^{on} motor neurons (arrowheads) are generated similarly in wild-type (**k, l**) and *Nkx2.2* mutant (**m, n**) embryos. **o, p, r, s**, Serotonin (5HT) expression in e11.5 (**o, r**) and e12.5 (**p, s**) wild-type (**o, p**) and *Nkx2.2* mutant (**r, s**) embryos. From e11.5, serotonin neurons are generated from *Nkx2.2*^{on} progenitors and are subsequently located adjacent to the floorplate (**p**). In *Nkx2.2* mutant embryos, the generation of these serotonin neurons is virtually eliminated. At e12.5, *Nkx2.9* expression is markedly downregulated in wild-type (**q**) and mutant (**t**) embryos. Serotonergic neurons of the dorsal raphe nuclei and dopaminergic neurons of the A9 (substantia nigra) and A10 (ventral tegmental area) nuclei are generated in normal numbers and positions in *Nkx2.2* mutant embryos. These monoaminergic neurons do not derive from *Nkx2.2*^{on} progenitor cells (data not shown).

zone²². The aberrant ventral position of motor neurons in *Nkx2.2* mutants appears therefore to be a consequence of their generation from 'x'-domain progenitors and not from ventrally directed migration. Moreover, most ventral somatic motor neurons are now generated from progenitor cells that do not express *Pax6* (Fig. 4c, i). This result provides further evidence that *Pax6* is not required directly for specification of somatic motor neuron fate but acts through repression of *Nkx2.2*. Together, these observations indicate that *Nkx2.2* expression in 'x'-domain progenitors suppresses their potential for somatic motor neuron generation (Fig. 4f, l).

In the hindbrain, 'x'-domain progenitors give rise to visceral motor neurons rather than to V3 neurons, but somatic motor neurons still derive from *Pax6*^{on} progenitors^{6,19}. The switch from V3 neuron to somatic motor neuron generation in the spinal cord of *Nkx2.2* mutant embryos suggested that 'x'-domain progenitors at hindbrain levels might undergo a corresponding ventral-to-dorsal switch, and generate somatic rather than visceral motor neurons. However, we found that the generation of somatic and visceral motor neurons in the hindbrain of *Nkx2.2* mutants was unchanged (Fig. 5b, c, g, h, k–n). In addition, the later pattern of migration of visceral (vagal) motor neurons (Fig. 5e, j) and the trajectory of the axons of hypoglossal and vagal motor neurons was similar in *Nkx2.2* mutant and wild-type embryos (data not shown). Thus, in the

hindbrain, *Nkx2.2* activity is not required for the generation of visceral motor neurons, nor does the inactivation of *Nkx2.2* result in a ventral expansion of the domain of somatic motor neuron generation (Fig. 5d, i).

What might account for the absence of a switch in motor neuron identity in the hindbrain? In the spinal cord, the extinction of *Nkx2.9* by 'x'-domain progenitors is virtually complete by e11.0 (Fig. 1e, h). Within the hindbrain, however, *Nkx2.9* and *Nkx2.2* expression overlaps until e11.0 (Fig. 1k). The pattern of expression of *Nkx2.9* is similar in *Nkx2.2* mutant embryos from e9.5 to e11.0 (Fig. 1n), and the expression of *Nkx2.9* is expanded dorsally in a manner similar to that of *Nkx2.2* in mice lacking *Pax6* (Fig. 1l, o). As *Nkx2.9* is closely related to *Nkx2.2* (ref. 13), it is likely that the normal generation of somatic and visceral motor neurons in the hindbrain of *Nkx2.2* mutants reflects the overlap in expression and redundant activity of *Nkx2.9*.

In the hindbrain, the period of visceral motor neuron generation persists until ~e11.0 (refs 6, 23) but later 'x'-domain progenitors appear to generate other cell types, notably serotonergic neurons²⁴. These neurons occupy a ventromedial position, adjacent to the floorplate (Fig. 5o, p). Although the expression of *Nkx2.2* persists during the period over which these serotonergic neurons are generated (Fig. 5o), by this time the expression of *Nkx2.9* is markedly downregulated (Fig. 5q). These observations raise the

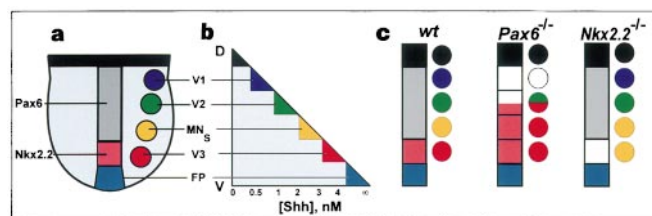


Figure 6 Distinct roles for Nkx2.2 and Pax6 in the control of ventral neuronal identity in response to graded Sonic hedgehog signalling. Progenitor cell domains, neuronal subtypes, their control by graded Shh signalling and the changes in cell identity that occur in *Pax6* and *Nkx2.2* mutants. **a**, Pax6 and Nkx2.2 progenitor cell domains and the dorsoventral position of generation of V1, V2 and V3 neurons and somatic motor neurons (MNs) in the ventral spinal cord. The position of the floor plate (FP) is shown and dorsal progenitor cells and neurons are indicated by dark grey shading. **b**, The relationship between Shh concentration and the generation of ventral interneurons, somatic motor neurons and floorplate cells, obtained from *in vitro* induction assays (modified from ref. 3). **c**, Changes in progenitor cell domains and neuronal subtype identity in *Pax6* and *Nkx2.2* mutants. WT: ventral pattern in wild-type embryos. Nkx2.2 is expressed in a ventral progenitor domain (the 'x' domain) adjacent to the floorplate. At spinal cord levels, these progenitors give rise to V3 neurons. Pax6 is expressed by all other ventral progenitor cells, in a dorsal-to-ventral gradient. The ventral-most population of Pax6^{on} progenitors gives rise to somatic motor neurons, whereas more dorsal Pax6^{on} progenitors give rise to V1 and V2 neurons. In *Pax6* mutants, there is a dorsal expansion in the domain of expression of Nkx2.2 (and *Nkx2.9*) in the position of generation of V3 neurons, and a reduction in somatic motor neurons and V2 neurons⁶. The generation of V1 neurons is lost in the absence of Pax6. In *Nkx2.2* mutants there is no change in Pax6 expression yet the fate of neurons is switched: 'x'-domain progenitors give rise to somatic motor neurons instead of V3 neurons. Shh signalling is unchanged in *Pax6* and *Nkx2.2* mutants, and the switch between somatic motor neuron and V3 neuron fate is a reflection of the pattern of Nkx2.2 expression. For details, see text.

possibility that the production of serotonergic neurons from 'x'-domain progenitors is affected in *Nkx2.2* mutants. Few, if any, serotonergic neurons were detected at the r2 level of the hindbrain in *Nkx2.2* mutants (Fig. 5r, s), despite the normal generation of trigeminal visceral motor neurons at the same axial level at an earlier stage (Fig. 5m). Thus, *Nkx2.2* function in hindbrain 'x'-domain progenitors is required for the development of serotonergic neurons. These results support the idea that *Nkx2.2* and *Nkx2.9* have overlapping functions in controlling the fate of 'x'-domain progenitors in the hindbrain. The apparent distinction in motor neuron phenotype observed in the spinal cord and hindbrain in *Nkx2.2*-mutant embryos may therefore reflect, in part, the differential timing of extinction of *Nkx2.9* at these two axial levels of the neural tube.

Together, these findings provide insight into the mechanisms by which progenitor cells interpret graded inductive signals to generate all-or-none distinctions in cell fate. Nkx2.2 has an essential role in interpreting graded Shh signalling and specifying the subtype identity of neurons generated from 'x'-domain progenitors (Fig. 6). The loss of *Nkx2.2* activity results in a 'ventral-to-dorsal' switch in progenitor fate, and the generation of neuronal progeny characteristic of cells that have been exposed to 2–3-fold less Shh activity⁶ (Fig. 6). Our results do not establish the sufficiency of Nkx2.2 in mediating the 'dorsal-to-ventral' switch in progenitor cell fate. Nevertheless in *Pax6* mutants, Nkx2.2 is expressed ectopically in more dorsal progenitors, and there is a concomitant dorsal-to-ventral switch in progenitor cell fate, leading to the generation of neuronal subtypes characteristic of exposure to 2–3-fold higher Shh concentration⁶ (Fig. 6). A *Drosophila* Nkx2.2 homologue, *vnd*, controls neuronal identities in the central nervous system in a manner similar to that shown here for Nkx2.2 (refs 25, 26). Loss of *vnd* function leads to a ventral-to-dorsal switch in neuronal fate

and overexpression of *vnd* induces a complementary dorsal-to-ventral switch in neuronal identity²⁶. Thus, although distinct inductive signals initiate ventral patterning in the vertebrate and insect central nervous systems, the *Nkx2.2/vnd* homeobox genes appear to represent an evolutionarily conserved step in this patterning process. □

Methods

Immunohistochemistry. Proteins were localized as described^{6,19,20}. Nkx2.2 was detected with a monoclonal antibody (mAb) or rabbit antiserum raised against chick Nkx2.2 (ref. 6). The rabbit anti-Pax6 antibody was a gift from S. Wilson²⁷. Nkx6.1 was detected with a rabbit antiserum²⁸. β-Tubulin was detected with mAb TuJ1 (ref. 29). Phox2A/B was detected with a rabbit antiserum³⁰ provided by J. Brunet. Serotonin was detected with a mAb (Research Biochemicals); Isl1/2, Lim3, HB9, Shh and HNF3β antisera have been described^{6,19}.

Localization of mRNA. *In situ* hybridization was performed as described⁶ using digoxigenin-labelled antisense riboprobes to *Sim1*, *Ngn3*, *Nkx2.2* and *Nkx2.9*.

Explant cultures. For *in vitro* cultures, ventral neural tube/floorplate explants were isolated from 33–36-somite (e10.5) wild-type and *Nkx2.2* mutants. Explants were dissected bilaterally at the ventral midline. Each half was cultured for 24 h either under control conditions or in the presence of mAb anti-Shh IgG 5E1 (ref. 16).

Nkx2.2 mutation. A null mutation in the *Nkx2.2* gene was generated by eliminating the entire coding region⁹. Genotyping of embryos was carried out by Southern-blot analysis as described⁹.

Received 3 February; accepted 2 March 1999.

- Smith, J. C. Mesoderm-inducing factors in early vertebrate development. *EMBO J.* **12**, 4463–4470 (1993).
- Gurdon, J. B., Dyson, S. & St Johnston, D. Cells perception of position in a concentration gradient. *Cell* **95**, 159–162 (1998).
- Ericson, J., Briscoe, J., Rashbass, P., van Heyningen, V. & Jessell, T. M. Graded sonic hedgehog signaling and the specification of cell fate in the ventral neural tube. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 451–466 (1997).
- Marti, E., Bumcrot, D. A., Takada, R. & McMahon, A. P. Requirement of 19K form of Sonic hedgehog for induction of distinct ventral cell types in CNS explants. *Nature* **375**, 322–325 (1995).
- Chiang, C. et al. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**, 407–413 (1996).
- Ericson, J. et al. Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell* **90**, 169–180 (1997).
- Barth, K. A. & Wilson, S. W. Expression of zebrafish *nkx2.2* is influenced by sonic hedgehog/vertebrate hedgehog-1 and demarcates a zone of neuronal differentiation in the embryonic forebrain. *Development* **121**, 1755–1768 (1995).
- Osumi, N. et al. Pax6 is involved in the specification of hindbrain motor neuron subtype. *Development* **124**, 2961–2972 (1997).
- Sussel, L. et al. Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development* **125**, 2213–2221 (1998).
- Ruiz i Altaba, A., Prezioso, V. R., Darnell, J. E. & Jessell, T. M. Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing centers: the dorsal lip/node, notochord and floor plate. *Mech. Dev.* **44**, 91–108 (1993).
- Hartigan, D. J. & Rubenstein, J. L. R. The cDNA sequence of murine Nkx-2.2. *Gene* **168**, 271–272 (1996).
- Shimamura, K., Hartigan, D. J., Martinez, S., Puelles, L. & Rubenstein, J. L. R. Longitudinal organization of the anterior neural plate and neural tube. *Development* **121**, 3923–3933 (1995).
- Pabst, O., Herbrand, H. & Arnold, H. H. Nkx2-9 is a novel homeobox transcription factor which demarcates ventral domains in the developing mouse CNS. *Mech. Dev.* **73**, 85–93 (1998).
- Yamada, T., Placzek, M., Tanaka, H., Dodd, J. & Jessell, T. M. Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. *Cell* **64**, 635–647 (1991).
- Qiu, M., Shimamura, K., Sussel, L., Chen, S. & Rubenstein, J. L. R. Control of anteroposterior and dorsoventral domains of Nkx-6.1 gene expression relative to other Nkx genes during vertebrate CNS development. *Mech. Dev.* **72**, 77–88 (1998).
- Ericson, J., Morton, S., Kawakami, A., Roelink, H. & Jessell, T. M. Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity. *Cell* **87**, 661–673 (1996).
- Fan, C. M. et al. Expression patterns of two murine homologs of *Drosophila* single-minded suggest possible roles in embryonic patterning and in the pathogenesis of Down syndrome. *Mol. Cell. Neurosci.* **7**, 519–535 (1996).
- Sommer, L., Ma, Q. & Anderson, D. J. Neurogenins, a novel family of atonal-related bHLH transcription factors, are putative mammalian neuronal determination genes that reveal progenitor cell heterogeneity in the developing CNS and PNS. *Mol. Cell. Neurosci.* **8**, 221–241 (1996).
- Tanabe, Y., William, C. M. & Jessell, T. M. Specification of motor neuron identity by the MNR2 homeodomain protein. *Cell* **95**, 67–80 (1998).
- Tsuchida, T. et al. Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell* **79**, 957–970 (1994).
- Sharma, K. et al. LIM homeodomain factors Lhx3 and Lhx4 assign subtype identities for motor neurons. *Cell* **95**, 817–828 (1998).
- Leber, S. M. & Sanes, J. R. Migratory paths of neurons and glia in the embryonic chick spinal cord. *J. Neurosci.* **15**, 1236–1248 (1995).
- Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A. & Krumlauf, R. Altered segmental identity and abnormal migration of motor neurons in mice lacking Hoxb-1. *Nature* **384**, 630–634 (1996).
- Lauder, J. & Bloom, F. Ontogeny of monoamine neurons in the locus coeruleus, Raphe nuclei and substantia nigra of the rat. I. Cell differentiation. *J. Comp. Neurol.* **155**, 469–481 (1974).

25. Jimenez, F. *et al.* *vnd*, a gene required for early neurogenesis of *Drosophila*, encodes a homeodomain protein. *EMBO J.* **14**, 3487–3495 (1995).
26. MacDonald, J. A. *et al.* Dorso-ventral patterning in the *Drosophila* CNS: The *vnd* homeobox gene specifies ventral column identity. *Genes Dev.* **12**, 3603–3612 (1998).
27. MacDonald, R. *et al.* The Pax protein *Noi* is required for commissural axon pathway formation in the rostral forebrain. *Development* **121**, 2397–2408 (1995).
28. Jensen, J., Serup, P., Karlsson, C., Nielsen, T. F. & Madsen, O. D. mRNA profiling of rat islet tumors reveals *nkx 6.1* as a beta-cell-specific homeodomain transcription factor. *J. Biol. Chem.* **271**, 18749–18758 (1996).
29. Geisert, E. E. & Frankfurter, A. The neuronal response to injury as visualized by immunostaining of class III beta-tubulin in the rat. *Neurosci. Lett.* **102**, 137–141 (1989).
30. Pattyn, A., Morin, X., Cremer, H., Coridis, C. & Brunet, J. F. Expression and interactions of the two closely related homeobox genes *Phox2a* and *Phox2b* during neurogenesis. *Development* **124**, 4065–4075 (1997).

Acknowledgements. We thank P. Rashbass and V. van Heyningen for *Sey* mice; D. Anderson, J. F. Brunet, H. Edlund, C. Goridis, A. Frankfurter, C. William and S. Wilson for reagents; C. Doe for results on *vnd*; R. Axel, K. Lee and G. Struhl for comments on the manuscript; and K. MacArthur for help in preparing the manuscript. This work was supported by grants to T.M.J. from the NIH and to J.L.R.R. from Nina Ireland, NARSAD and NIMH; by an HFSP fellowship (J.B.); by the Swedish Institute, the Swedish Foundation for Strategic Research, and the Swedish National Research Council (J.E.); by the Bank of America, Scottish Rite and NIH, NRSA (L.S.); and by the Danish National Research Foundation (P.S.). J.B. and J.E. are associates and T.M.J. is an investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to J.L.R.R. (e-mail: jlrr@cgl.ucsf.edu) or T.M.J. (e-mail: tmj1@columbia.edu).

Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms

Gijsbertus T. J. van der Horst*, Manja Muijtens*, Kumiko Kobayashi†, Riya Takano†, Shin-ichiro Kanno†, Masashi Takao†, Jan de Wit*, Anton Verkerk*, Andre P. M. Eker*, Dik van Leenen‡, Ruud Buijs§, Dirk Bootsma*, Jan H. J. Hoeijmakers* & Akira Yasui†

* MGC, Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

† Department of Molecular Genetics, Institute of Development, Aging and Cancer, Tohoku University, 980-8575 Sendai, Japan

‡ MGC, Department of Clinical Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

§ Netherlands Institute for Brain Research, Meibergdreef 33, 1105 AZ Amsterdam, The Netherlands

Many biochemical, physiological and behavioural processes show circadian rhythms which are generated by an internal time-keeping mechanism referred to as the biological clock. According to rapidly developing models, the core oscillator driving this clock is composed of an autoregulatory transcription–(post) translation-based feedback loop involving a set of ‘clock’ genes^{1–6}. Molecular clocks do not oscillate with an exact 24-hour rhythmicity but are entrained to solar day/night rhythms by light. The mammalian proteins Cry1 and Cry2, which are members of the family of plant blue-light receptors (cryptochromes) and photolyases, have been proposed as candidate light receptors for photoentrainment of the biological clock^{7–10}. Here we show that mice lacking the Cry1 or Cry2 protein display accelerated and delayed free-running periodicity of locomotor activity, respectively. Strikingly, in the absence of both proteins, an instantaneous and complete loss of free-running rhythmicity is observed. This suggests that, in addition to a possible photoreceptor and antagonistic clock-adjusting function, both proteins are essential for the maintenance of circadian rhythmicity.

We were interested in identifying mammalian homologues of the DNA-repair enzyme photolyase, a protein that undoes ultraviolet-induced DNA damage in a single-step process (photoreactivation) requiring light energy captured by blue-light-collecting chromophores^{11,12}. In this search, we and others have cloned two genes with strong homology to class I photolyases of lower species^{7–10}. In addition to the photolyase core domain, the gene products appeared to contain a carboxy-terminal extension also

found in plant blue-light receptors (cryptochromes), for which the mammalian photolyase-like genes were designated *cry1* and *cry2*. Plant Cry proteins mediate light-dependent processes such as phototropism, growth and flowering^{13–15}. Since placental mammals as well as endogenous or recombinant mammalian Cry proteins lack clearly detectable photoreactivating activity, the mammalian Cry proteins may act as photoreceptors rather than photolyases^{7,9,10}. The biological ‘master’ clock in the suprachiasmatic nucleus (SCN) of the brain controls many physiological processes, from body temperature to the sleep–wake cycle. A major question in mammalian chronobiology is how the clock is entrained to solar time, thereby keeping an organism in an exact 24-h rhythm. The absence of photoentrainment in eye-less rodents indicates that the light receptors feeding into the SCN circadian system must reside in the eye^{16,17}, but the process does not seem to depend on retinal photoreceptor cells and their visual pigments, as Retinal-degenerate (*Rd*) mice show a normal circadian response to light^{17,18}. Since mammalian *cry* genes are specifically expressed in the ganglion and inner nuclear layer of the retina, the Cry1 and Cry2 proteins are possible candidates for circadian photoreceptors¹⁹.

To explore the biological function of mammalian Cry1 and Cry2, we have generated *cry1* and *cry2* mutant mice through gene targeting in embryonic stem cells (Fig. 1). Analysis of the transcriptional status of the targeted *cry1* and *cry2* genes, using the reverse transcription-long-range polymerase chain reaction, revealed no detectable transcripts in the corresponding knockout animals, thus demonstrating that we have created null-mutant mice. Targeted *cry1* and *cry2* alleles both segregate at expected mendelian ratios, indicating that the absence of either Cry1 or Cry2 does not interfere with embryonic development. Moreover, *cry1* and *cry2* mutant mice are completely healthy and show no overt phenotype (the oldest animals are now 14 and 7 months, respectively). We analysed the possible role for Cry proteins in the biological clock by measuring the circadian wheel-running behaviour of *cry*-knockout mice under normal light/dark (LD) cycles and in constant darkness (dark/dark; DD). We made two unexpected observations. First, compared with wild-type mice which, when subjected to DD conditioning, have a free-running rhythm close to 24 hours ($\tau = 23.77 \pm 0.07$ h ($n = 14$)), the internal clock of *cry1* mutants runs significantly faster ($\tau = 22.51 \pm 0.06$ h ($n = 9$); $P < 0.00001$) (Fig. 2a, b). In contrast, *cry2* mutants exhibit a clear increase in period length ($\tau = 24.63 \pm 0.06$ h ($n = 5$); $P < 0.00001$) (Fig. 2c). Heterozygous animals showed wheel-running patterns comparable to wild-type mice, and there were no clear sex- or age-related differences (data not shown). These findings suggest that Cry1 and Cry2 antagonistically modulate the period length of the clock. Second, under LD conditions, both mutants show a circadian periodicity of 24 h (Fig. 2a–c), suggesting that a deficiency in either *cry1* or *cry2* does not produce a detectable loss of light entrainment of locomotor activity. However, as *cry1* mice still contain a functional Cry2 protein which may (partly) take over the function of Cry1 (and vice versa), functional redundancy may blur the phenotypic outcome. Thus, it was of interest to examine double-mutant mice.

Like *cry*-single-knockout mice, double-mutant animals are viable and show no gross phenotypic abnormalities (by 5 months old). Unexpectedly, these mice ($n = 8$) still display an essentially 24-h circadian rhythm under LD conditions (Fig. 2f, upper part). This suggests that, despite the absence of both Cry proteins, the biological clock—as reflected by locomotor activity—may still receive a light input. However, when double-mutant mice are shifted to a DD regime, they show a striking instantaneous and complete circadian arrhythmicity (Fig. 2f). This indicates that there is no internal circadian clock running with any significant momentum, although we do not exclude the possibility that there is still an ultradian component. Note also that the instantaneous arrhythmicity in double-mutant mice differs from any clock mutant analysed thus far. This includes the only mouse ‘clock’ mutant (*clock*) described to

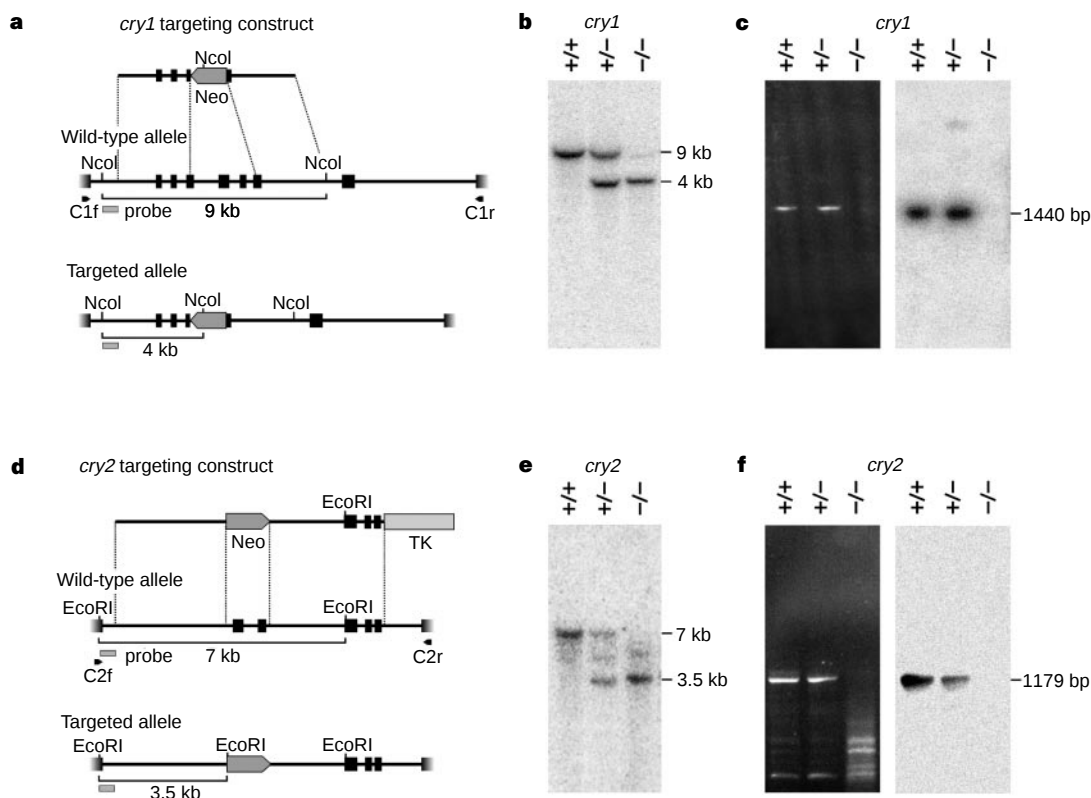


Figure 1 Targeted disruption of the *cry* genes and generation of Cry-deficient mice. **a**, Physical map of the wild-type *cry1* locus, the targeting construct and the disrupted *cry1* locus. Exons are indicated by black filled boxes. Note that the use of PCR-derived genomic DNA does not allow proper exon numbering. The probe used for screening homologous recombinants and genotyping mice, localized external to the construct, is represented by a grey box. Primers used for RNA analysis by RT-long-range PCR are depicted as black arrowheads. **b**, Southern blot analysis of tail DNA of *cry1*^{+/+}, *cry1*^{+/-} and *cry1*^{-/-} mice, obtained from heterozygote intercrosses. In *Nco*I-digested DNA, the *cry1* probe hybridizes to a 9-kilobase (kb) fragment of the wild-type *cry1* locus and a 4-kb fragment of the targeted locus. **c**, Expression of the wild-type and targeted *cry1* locus. The left panel shows ethidium-bromide-stained RT-long-range PCR products obtained from *cry1*^{+/+},

cry1^{+/-} and *cry1*^{-/-} mouse embryonic fibroblasts using primer set C1f/C1r. The absence of the 1,440-bp *cry1*-specific PCR product in *cry1*^{-/-} animals is confirmed by Southern blot analysis using the complete *cry1* cDNA as a probe (right panel). **d**, Physical map of the wild-type *cry2* locus, the targeting construct and the disrupted *cry2* locus. The use of symbols is explained in **a**. **e**, Southern blot analysis of tail DNA of *cry2*^{+/+}, *cry2*^{+/-} and *cry2*^{-/-} mice, obtained from heterozygote intercrosses. The *cry2* probe hybridizes to 7- and 3.5-kb *Eco*RI fragments of the wild-type and disrupted *cry2* locus, respectively. **f**, Expression of the wild-type and targeted *cry2* locus. Left panel, ethidium-bromide-stained RT-long-range PCR products obtained from kidney RNA using primer set C2f/C2r; right panel, the corresponding Southern blot probed with *cry2* cDNA and used to discriminate between specific and background fragments.

date, which shows a more gradual loss of periodicity^{20,21}. Interestingly, mice ($n = 4$) with only one functional *cry2* allele out of the four *cry* gene copies initially display a free-running rhythm even shorter than *cry1*-knockout mice, which gradually progresses into arrhythmicity (Fig. 2e). Apparently, these mice possess a clock, but one dose of *cry2* can only keep it running for a limited number of cycles in the absence of light. On transfer into LD again these mice regain their original clock, showing that the gradual loss of rhythmicity in DD is reversible (data not shown). This demonstrates a direct involvement of Cry2 in maintaining the clock. On the other hand, in the presence of one allele of *cry1*, mice ($n = 4$) show rhythmic activity but the periodicity is intermediate to that of *cry1* and *cry2* single-mutant mice (Fig. 2d). Mice heterozygous for both *cry1* and *cry2* show wild-type DD wheel-running patterns (data not shown). Thus, our results demonstrate that the Cry proteins are involved in maintaining period length and circadian rhythmicity, and that a critical balance between Cry1 and Cry2 is required for proper clock functioning.

As the Cry proteins were expected to function as circadian photoreceptors, it was surprising to see an apparent effect of light on the running-wheel behaviour of totally Cry-deficient animals (Fig. 2f, LD part). Therefore, we analysed the behaviour of wild-type and double-mutant mice under changing light conditions. Photoentrained wild-type mice ($n = 4$), when shifted from the normal LD

(12:12 h) regime to a different LD regime (6:18 h), maintain a virtually normal period of activity that gradually shifts towards the new 'lights-off' set-point according to the free-running rhythmicity (Fig. 3a). In sharp contrast, *cry* double-mutant mice ($n = 4$) abruptly adapt to the new situation by starting wheel running as soon as the light is off and by expanding their period of fragmented activity to 18 h (Fig. 3b). When given daily random blocks of 8 h of light, wild-type mice ($n = 5$) try to adjust their biological clock, whereas double-mutant *cry* mice ($n = 5$) maintain arrhythmic wheel-running behaviour that is only interrupted by light (Fig. 3d–f). Moreover, the running activity of wild-type mice seems also to be transiently and completely suppressed by light (see Fig. 3d, arrow). These observations suggest that light has a dominant influence on the behaviour of (nocturnal) mice, preventing them from using the running wheel when the light is on. Thus, our findings support the idea that under a normal light regime (LD 12:12 h) a dominant effect of light on the running behaviour masks the defective biological clock in totally *cry*-deficient mice, apparent from their behaviour in DD (see also Fig. 3c). We are currently generating rod-less *cry* mutant mice to study in detail the photoreceptor function of the *cry* proteins in the absence of the visual light-perceptive system.

How do Cry1 and Cry2 fit into the rapidly advancing model of the molecular mechanism of the clock? Recently, behavioural analysis of

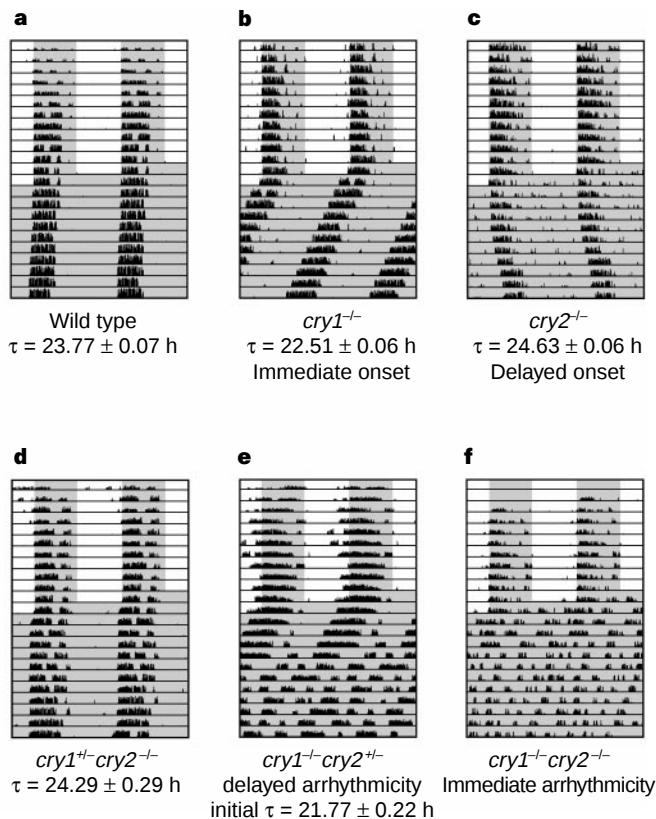


Figure 2 Circadian phenotype of *cry* mutant mice. Representative examples are shown of voluntary wheel-running records of wild-type ($n = 14$), *cry1*^{-/-} ($n = 9$), *cry2*^{-/-} ($n = 5$), *cry1*^{+/-}*cry2*^{-/-} ($n = 4$), *cry1*^{-/-}*cry2*^{+/-} ($n = 4$) and *cry1*^{-/-}*cry2*^{-/-} ($n = 8$) mutant mice under LD (12:12 h) and DD conditions. Daily activity patterns are double plotted (for example, day 1 + 2, day 2 + 3, day 3 + 4) on consecutive lines. Periods of darkness are indicated by a grey background. Free-running periods (τ ± s.e.m.) were calculated as described in the Methods section.

cry2-knockout mice has revealed altered photoresponses but persistent circadian rhythmicity in DD²², which is consistent with our findings for both *cry* mutants. This is consistent with the idea that the Cry proteins are circadian photoreceptors. However, our observations with *cry* double-mutants and with mice carrying one active *cry2* allele disclose a second important function of Cry proteins in mammals, namely, involvement in maintaining rhythmicity in constant darkness. This indicates that the Cry proteins not only function in the light-input pathway but are also clock components. As such, expression of the mammalian *cry* genes resembles that of the clock genes (see ref. 1 and refs therein, and refs 23, 24) in that they are not only expressed in the retina and SCN but also in all other organs and tissues analysed^{8,10,18}. Moreover, the *cry1* gene is rhythmically expressed in the mouse SCN even in the absence of Cry2²². Using mutant and overexpressing flies, the recently discovered *drosophila* *cry* gene product has also been shown to act as a photoreceptor^{25,26}. In *cry* mutant flies the rhythmically expressed clock genes *per* and *tim* show flat expression levels, except in a subset of lateral neurons. We do not exclude the possibility that, analogous to the mammalian situation, *Drosophila* also contains a second *cry* gene, which may explain the persisting rhythmic behaviour of the mutant flies.

Future research will focus on the precise roles of the mammalian *cry* genes in the master 'clock' in the SCN and the autonomous clock in every cell. In particular, it will be important to know how these proteins are involved in regulating the expression of known clock genes, or in light-dependent stabilization of (clock) proteins. Finally, one interesting aspect is the subcellular localization of the

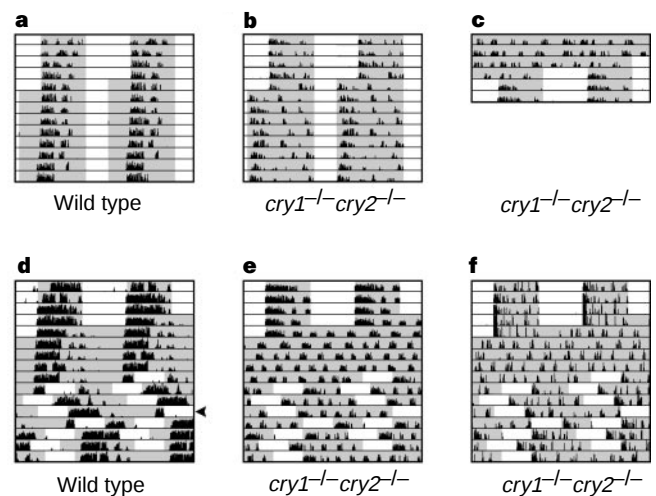


Figure 3 Circadian phenotype of *cry1* double-mutant mice under changing light conditions. Wild-type ($n = 4$) and *cry* double-mutant mice ($n = 4$) were maintained under normal LD (12:12 h) and subsequently exposed to a short day (LD, 6:18 h) light regime (a, b). In addition, animals were maintained under continuous darkness (DD) and shifted to a normal LD (12:12 h) protocol (c, representative example of 6 *cry1*^{-/-}*cry2*^{-/-} mice) or a light regime where they encountered random daily light blocks of 8 h (d-f, representative examples of 5 wild-type and 5 *cry1*^{-/-}*cry2*^{-/-} mice). The arrowhead in d illustrates the suppressing effect of light on wheel-running behaviour of a wild-type mouse. Periods of darkness are indicated by a grey background.

Cry proteins. Previously, we have shown that, in mouse liver cells and cultured human fibroblasts, Cry2 resides in both the nucleus and mitochondria, whereas all detectable Cry1 appears to be in mitochondria¹⁰. These findings suggest that mitochondria may play an important role in controlling the biological clock. This may not be without precedent, as inhibitors of mitochondrial function can induce a phase-shift in the circadian rhythm of *Neurospora*^{27,28}. Future research must focus on the intriguing connection between this organelle and the biological clock. □

Methods

Generation of *cry1* and *cry2* targeting constructs. Isogenic mouse genomic DNA was obtained by amplification of Ola129-derived E14 ES cell DNA (Takara, LA-PCR; E14 line was provided by A. Berns, Netherlands Cancer Institute) using primer sets designed on the basis of the mouse *cry1* and *cry2* complementary DNA sequence (GenBank accession nos 000777 and 003433, respectively¹⁰). The construction of the *Neo*-selectable *cry1* and *cry2* targeting vectors (see Fig. 1a, d) used to delete coding sequences corresponding to base pairs (bp) 730–1,479 and 397–810 of the respective cDNA sequences and isolation of genomic probes will be described in detail elsewhere. Note that the use of PCR-derived genomic DNA does not allow appropriate exon numbering.

Disruption of *cry1* and *cry2* in mouse ES cells. ES cell line E14 was maintained on gelatine-coated dishes in 60% BRL conditioned DMEM/40% fresh DMEM medium, supplemented with 15% fetal calf serum, 0.1 mM non-essential amino acids, 2 mM glutamine, 50 $\mu\text{g ml}^{-1}$ penicillin and streptomycin, 1,000 U ml⁻¹ leukaemia inhibitory factor (all components purchased from Gibco) and 0.1 mM 2-mercaptoethanol. Linearized targeting vector DNA (15 μg) was transfected into E14 cells (1×10^7 cells in 400 μl PBS) by electroporation for 10 ms at 1,200 μF and 117 V, using a Progenitor II Gene Pulser (Hoeffer). Electroporated cells were reseeded onto 10-cm dishes and subjected to neomycin selection by addition of 200 $\mu\text{g ml}^{-1}$ G418 (Geneticin, Gibco) the following day. In the case of *cry2*, counterselection against randomly integrated DNA was obtained by including 0.2 μM flauridine in the selection medium. After nine days colonies were randomly picked and expanded in 24-well dishes. Duplicate dishes were used for cryopreservation and genotyping (Southern blot analysis), respectively. Both experiments produced targeting frequencies of 5%.

Generation of *cry1*- and *cry2*-knockout mice. Gene-targeted ES cells, checked for proper chromosome composition by karyotyping, were injected into C57BL/6 blastocysts by standard procedures²⁹. Chimaeric male mice were mated with C57BL/6 females and transmission of E14-derived germ cells was identified by an agouti coat colour in the offspring. Heterozygous male and female mice were interbred to generate *cry1* and *cry2* and *cry1/cry2* (double) knockout mice.

DNA and RNA analysis. Genotyping of ES cell, mouse embryonic fibroblast (MEF) or tail DNA was performed by Southern blot analysis. For *cry1*, *NcoI*-digested DNA was probed with a 560-bp *NcoI*-*XbaI* fragment. For *cry2*, *EcoRI*-digested DNA was probed with a 519-bp PCR fragment flanking the 5' border of the construct. Both probes were located outside the targeting construct. RNA was examined by RT-long-range PCR (Takara) analysis using primer sets C1F (5'-CGCATTTCACATACACTGTATGACCTGGACAA-3')/C1R (5'-TTACTGCTCAGCTGCTGGGACTTTG-3') and C2F (5'-GTGCACTGGTTCGGAAGGG-3')/C2R (5'-AGCACTGCAGGACAGCCACA-3') for *cry1* and *cry2*, respectively. Randomly primed cDNA was synthesized from RNA isolated with a Rneasy kit (Qiagen). The specificity of the PCR products was confirmed by Southern blot analysis using complete cDNA probes.

Monitoring of locomotor-activity rhythm. Mice of either sex and of different ages (8–20 weeks) were used. They were individually housed in a light-proof chamber in cages (30 × 45 cm) equipped with a running wheel (11 cm in diameter) and a magnetic sensor system to detect wheel rotations³⁰. Animals were maintained in a cycle of 12 h light (150 lux) and 12 h complete darkness (LD) or in continuous complete darkness (DD) in constant ambient temperature (21 ± 0.5 °C) with water and food available *ad libitum*. Voluntary wheel running (wheel turns per unit of time) was continuously recorded by an online IBM computer using a modified version of the Rodent Activity Meter program³⁰. Activity records were plotted as actograms and the period of locomotor activity was determined from the slope of an eye-fitted straight line through the daily activity onset. The reproducibility of this method was demonstrated by the minimal difference in period estimations (<5 min) when performed by three independent investigators. Unpaired Student's *t*-tests were used to make statistical comparisons between the different genotypes.

Received 27 November 1998; accepted 8 March 1999.

- Whitmore, D., Sassone-Corsi, P. & Foulkes, N. S. PASTing together the mammalian clock. *Curr. Opin. Neurobiol.* **8**, 635–641 (1998).
- Wilsbacher, L. D. & Takahashi, J. S. Circadian rhythms: molecular basis of the clock. *Curr. Opin. Genet. Dev.* **8**, 595–602 (1998).
- Reppert, S. M. A clockwork explosion! *Neuron* **21**, 1–4 (1998).
- Dunlap, J. An end in the beginning. *Science* **280**, 548–549 (1998).
- Sassone-Corsi, P. Molecular clocks: mastering time by gene regulation. *Nature* **392**, 871–874 (1998).
- Schibler, U. New cogwheels in the clockwork. *Nature* **393**, 620–621 (1998).
- Todo, T. *et al.* Similarity among the *Drosophila* (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family. *Science* **272**, 109–112 (1996).
- van der Spek, P. J. *et al.* Cloning, tissue expression, and mapping of a human photolyase homolog with similarity to plant blue-light receptors. *Genomics* **37**, 177–182 (1996).
- Hsu, D. *et al.* Putative human blue-light photoreceptors hCRY1 and hCRY2 are flavoproteins. *Biochemistry* **35**, 13871–13877 (1996).
- Kobayashi, K. *et al.* Characterization of photolyase/blue-light receptor homologs in mouse and human cells. *Nucleic Acids Res.* **26**, 5086–5092 (1998).
- Sancar, A. Structure and function of DNA photolyase. *Biochemistry* **33**, 2–9 (1994).
- Yasui, A. & Eker, A. P. M. in *DNA Damage and Repair, Vol. 2: Repair in Higher Eukaryotes* (eds Nickoloff, J. A. & Hoekstra, M. E.) 9–31 (Humana, Totowa, NJ, 1998).
- Ahmad, M., Jarillo, J. A. & Cashmore, A. R. Chimeric proteins between *cry1* and *cry2 Arabidopsis* blue light photoreceptors indicate overlapping functions and varying protein stability. *Plant Cell* **10**, 197–207 (1998).
- Guo, H., Yang, H., Mockler, T. C. & Lin, C. Regulation of flowering time by *Arabidopsis* photoreceptors. *Science* **279**, 1360–1363 (1998).
- Suárez-López, P. & Coupland, G. Plants see the blue-light. *Science* **279**, 1323–1324 (1998).
- Nelson, R. J. & Zucker, I. Absence of extraocular photoreception in diurnal and nocturnal rodents exposed to direct sunlight. *Comp. Biochem. Physiol. A* **69**, 145–148 (1981).
- Foster, R. G. Shedding light on the biological clock. *Neuron* **20**, 829–832 (1998).
- Foster, R. G. *et al.* Circadian photoreception in the retinally degenerate mouse (rd/rd). *J. Comp. Physiol. A* **619**, 39–50 (1991).
- Miyamoto, Y. & Sancar, A. Vitamin B₂-based blue-light photoreceptors in the retinohypothalamic tract as the photoactive pigments for setting the circadian clock in mammals. *Proc. Natl Acad. Sci. USA* **95**, 6097–6102 (1998).
- Kings, D. P. *et al.* Positional cloning of the mouse circadian clock gene. *Cell* **89**, 641–653 (1997).
- Antoch, M. P. *et al.* Functional identification of the mouse circadian clock gene by transgenic BAC rescue. *Cell* **89**, 655–667 (1997).
- Trescher, R. J. Role of mouse cryptochrome blue-light photoreceptor in circadian photoresponses. *Science* **282**, 1490–1494 (1998).
- Sangoram, A. M. *et al.* Mammalian circadian autoregulatory loop: A Timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription. *Neuron* **21**, 1101–1113 (1998).
- Zylka, M. J. *et al.* Molecular analysis of mammalian timeless. *Neuron* **21**, 1115–1122 (1998).
- Emery, P., So, W. V., Kaneko, M., Hall, J. C. & Rosbash, M. CRY, a *Drosophila* clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. *Cell* **95**, 669–679 (1998).

- Stanewsky, R. *et al.* The *cry^b* mutation identifies cryptochrome as a circadian photoreceptor in *Drosophila*. *Cell* **95**, 681–692 (1998).
- Brody, S. Circadian rhythms in *Neurospora crassa*: the role of mitochondria. *Chronobiol. Int.* **9**, 222–230 (1992).
- Millar, A. J. Circadian rhythms: PASSing time. *Current Biology* **7**, R474–R476 (1997).
- Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells. A Practical Approach* (ed. Joyner, A. L.) 113–151 (IRL, Oxford, 1987).
- Horton, R. M. Simple, inexpensive computerized rodent activity meters. *Biotechniques* **19**, 594–597 (1995).

Acknowledgements. We thank J. Miyazaki for help with construction of targeting vectors, and Y. Tsukahara and S. Okano for discussions on circadian rhythms. M. Kuit is acknowledged for photographic work. This work was supported in part by grants from the Dutch Cancer Society, the Association for International Cancer Research, Human Frontier Science, The Louis Jeantet Foundation and the Ministry of Education, Science, Sports and Culture of Japan, and by a Spinoza premium of the Dutch Scientific Organization NWO.

Correspondence and requests for materials should be addressed to J.H.J.H.

Direct control of the Forkhead transcription factor AFX by protein kinase B

Geert J. P. L. Kops, Nancy D. de Ruiter, Alida M. M. De Vries-Smits, David R. Powell*, Johannes L. Bos & Boudewijn M. Th. Burgering

Laboratory for Physiological Chemistry and Centre for Biomedical Genetics, University of Utrecht, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands

* Department of Pediatrics, Baylor College of Medicine, Houston, Texas 77030, USA

The phosphatidylinositol-3-OH-kinase (PI(3)K) effector protein kinase B (refs 1, 2) regulates certain insulin-responsive genes^{3,4}, but the transcription factors regulated by protein kinase B have yet to be identified. Genetic analysis in *Caenorhabditis elegans* has shown that the Forkhead transcription factor *daf-16* is regulated by a pathway consisting of insulin-receptor-like *daf-2* and PI(3)K-like *age-1* (refs 5–8). Here we show that protein kinase B phosphorylates AFX, a human orthologue of *daf-16* (refs 5, 6, 9), both *in vitro* and *in vivo*. Inhibition of endogenous PI(3)K and protein kinase B activity prevents protein kinase B-dependent phosphorylation of AFX and reveals residual protein kinase B-independent phosphorylation that requires Ras signalling towards the Ral GTPase. In addition, phosphorylation of AFX by protein kinase B inhibits its transcriptional activity. Together, these results delineate a pathway for PI(3)K-dependent signalling to the nucleus.

To test for potential regulation of the *daf-16* orthologue AFX (refs 5, 6, 9) in mammalian cells, we labelled A14 cells transiently expressing haemagglutinin-epitope-tagged AFX (HA-AFX) with ³²P-orthophosphate and treated them with insulin. HA-AFX underwent a rapid and sustained increase in phosphorylation following insulin treatment (Fig. 1a). Treatment of A14 cells with epidermal growth factor and of Rat1 cells with platelet-derived growth factor also increased phosphorylation of HA-AFX (data not shown). Phospho-amino-acid analysis of immunoprecipitated AFX revealed that AFX was phosphorylated on serine and threonine residues (Fig. 1b). To investigate whether PI(3)K and protein kinase B (PKB) were involved in insulin-induced phosphorylation of AFX, we co-expressed HA-AFX with active forms of PI(3)K or PKB. Constitutively active, but not inactive, forms of PI(3)K or PKB induced a strong increase in HA-AFX phosphorylation (Fig. 1c).

AFX contains three putative PKB phosphorylation sites¹⁰ (T28, S193 and S258; Fig. 2a) that are conserved between AFX and *daf-16*. As PKB activation *in vivo* was sufficient to increase HA-AFX phosphorylation, we investigated whether AFX could be phosphorylated by PKB *in vitro*. As shown in Fig. 2b, both immunoprecipitated HA-AFX (right panel) and a bacterially expressed fusion with glutathione-S-transferase (GST-AFX; left panel) were

phosphorylated by purified active baculo-PKB¹¹. Mutating any one of the three putative PKB phosphorylation sites did not produce a marked decrease in phosphorylation by PKB, suggesting that at least two sites were phosphorylated by PKB (Fig. 2b). Tryptic peptide mapping of *in vitro* phosphorylated wild-type HA-AFX revealed three reproducibly radiolabelled peptides designated 1, 2 and 3 (Fig. 2c). Peptides 1 and 2 contained S258 and S193, respectively because peptide maps of the *in vitro* phosphorylated S258A and S193A mutants no longer displayed ³²P-incorporation into the corresponding peptides (Fig. 2c). Susceptibility of peptide 1, but not 2, to secondary thermolysin digestion, manual Edman degradation (release of ³²P in the third cycle) and phospho-amino-acid analysis of eluted peptides 1 and 2 (serine phosphorylation only) supported this conclusion (data not shown). The phosphorylation of peptide 3 was not due to phosphorylation of T28, as the T28A mutation did not abolish phosphorylation of peptide 3. Therefore, PKB can phosphorylate proteins on residues not lying within the consensus sequence for PKB, albeit at a low stoichiometry (Fig. 2c). Substitution of either S193 or S258 to a threonine residue (S193T and S258T, respectively) proved that these two specific residues are phosphorylated by PKB, because these mutations conferred phosphorylation of a threonine residue on HA-AFX, whereas only serine residues were phosphorylated in wild-type HA-AFX (Fig. 2d). Peptide mapping of *in vivo* phosphorylated HA-AFX showed that the same three peptides are phosphorylated, as well as one additional peptide, designated peptide 4 (Fig. 2e). Phosphorylation of peptides 1 and 2 was induced when cells were treated with insulin or when active PKB (gagPKB) was co-expressed (Fig. 2e). Peptide map analysis of the S193A and S258A mutants showed that these two residues were also phosphorylated *in vivo* following insulin treatment (Fig. 2f). From these results, we conclude that PKB predominantly phosphorylates S193 and S258, both *in vitro* and *in vivo*.

Three observations suggest that there is an insulin-induced, PKB-independent pathway acting on AFX. First, *in vitro*, PKB induces only serine phosphorylation. Second, *in vivo*, one PKB-independent

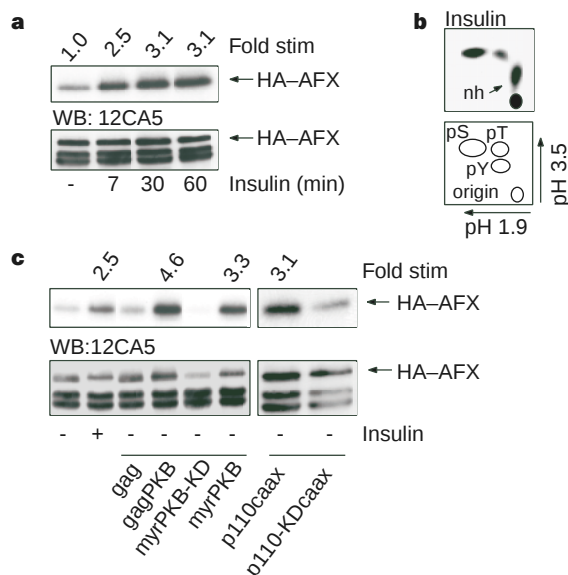


Figure 1 Insulin treatment and active forms of PI(3)K and PKB increase phosphorylation of HA-AFX *in vivo*. **a**, HA-AFX was immunoprecipitated from ³²P-orthophosphate labelled A14 cells left untreated (–) or treated with insulin for 7, 30 or 60 minutes. Following exposure to film, the blot was probed with 12CA5 monoclonal to ensure equal expression of HA-AFX in each lane (WB: 12CA5). Fold stim; fold increase in phosphorylation over control. **b**, ³²P-labelled HA-AFX (30 minute insulin time-point) was subjected to two-dimensional analysis²⁸ (nh: non-hydrolysed protein). Positions of phospho-amino-acids are as indicated. **c**, A14 cells were transfected with HA-AFX combined with the indicated cDNAs. Phosphorylation of HA-AFX was analysed as in **a**.

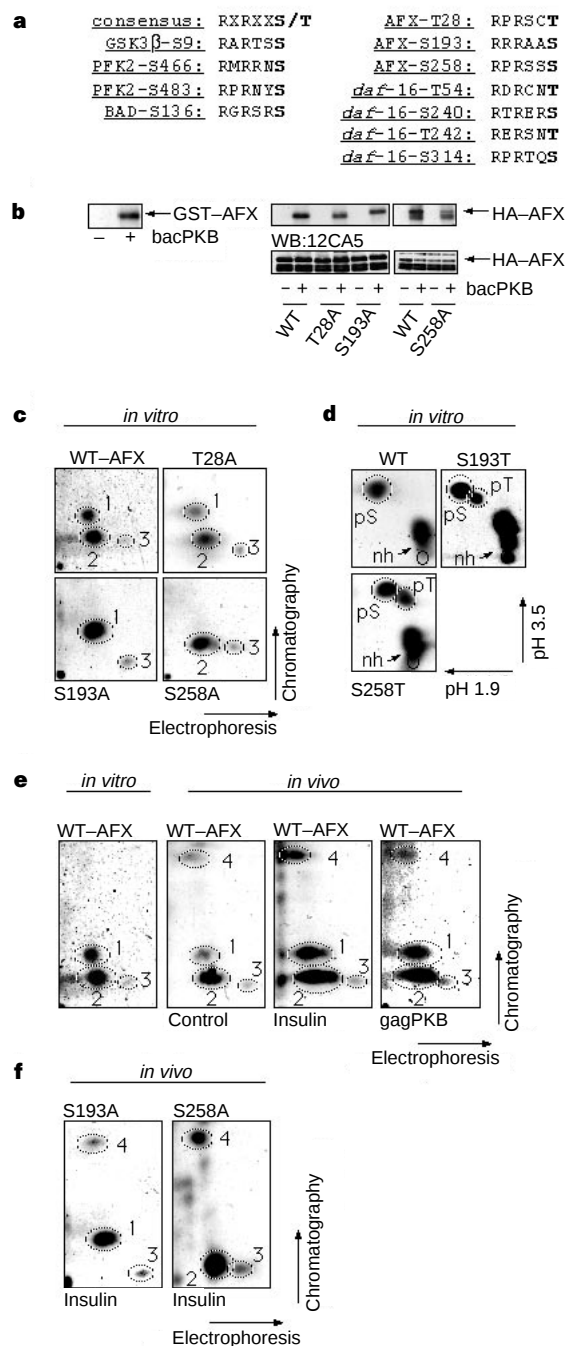


Figure 2 PKB phosphorylates HA-AFX *in vitro* and *in vivo*. **a**, Comparison of putative PKB phosphorylation sites in HA-AFX (T28, S193 and S258) and *daf-16* with known PKB sites (BAD, GSK-3 β and PFK2) and PKB consensus as determined by Alessi *et al.* (ref. 10). **b**, Immunoprecipitated wild-type HA-AFX and PKB-site mutants of HA-AFX (T28A, S193A and S258A) or purified bacterially expressed GST-AFX were incubated in the presence (+) or absence (–) of 2 μ l of purified active baculo-PKB (bacPKB). **c**, HA-AFX phosphorylated *in vitro* by PKB as in **b** was processed for two-dimensional phospho-peptide mapping as described²⁸. Positions of the three resolved peptides are indicated by numbers. **d**, HA-AFX, S193T and S258T phosphorylated *in vitro* by PKB as in **b** were processed for two-dimensional phospho-amino-acid analysis as in Fig. 1b. nh: non-hydrolyzed protein. **e**, HA-AFX phosphorylated *in vivo* following insulin treatment of A14 cells or by co-expression with gagPKB were isolated as in Fig. 1a, and analysed as described in **c**. **f**, S193A and S258A phosphorylated *in vivo* following insulin treatment of A14 were isolated as in Fig. 1a, and analysed as described in **c**.

peptide is phosphorylated (peptide 4; Fig. 2e and f). Finally, both complete inhibition of PI(3)K (and thus PKB, data not shown and ref. 1) by pretreatment of A14 cells with wortmannin¹² (Fig. 3a) or LY294002 (ref. 12; data not shown), and inhibition of endogenous PKB activity by expressing a dominant-negative PKB (PKBcaax)¹¹, resulted in a pronounced but incomplete inhibition of insulin-induced HA-AFX phosphorylation (Fig. 3a). Treatment of A14 cells with insulin activates many signalling cascades, including those elicited by activation of Ras¹³. Expression of an active mutant of Ras (RasV12)¹⁴ induced an insulin-independent phosphorylation of HA-AFX (Fig. 3b), in keeping with the ability of oncogenic Ras to induce PI(3)K-dependent signalling¹⁵. Furthermore, inhibiting Ras activation by expressing dominant-negative Ras (RasN17)¹⁴, which does not inhibit PKB activation (data not shown and ref. 1), partially inhibited HA-AFX phosphorylation induced by insulin (Fig. 3b). Simultaneous blocking of Ras and PI(3)K using RasN17 and wortmannin completely blocked insulin-induced phosphorylation of HA-AFX (Fig. 3c). As well as PI(3)K, Ras activates the

RalGEF/Ral and Raf/MEK/MAPK pathways^{16,17}. Expression of active Rlf (Rlfcaax), a guanylyl-exchange factor for Ral¹⁸, resulted in phosphorylation of HA-AFX, whereas active Raf¹⁹ or a catalytically inactive form of Rlf had no effect (Fig. 3d). Dominant-negative RalN28 (ref. 20) completely abolished the Rlfcaax-induced phosphorylation of HA-AFX, whereas continuous inhibition of PI(3)K using wortmannin had no effect (Fig. 3e). Combined inhibition of both Ral and PI(3)K using RalN28 and wortmannin completely blocked insulin-induced phosphorylation of HA-AFX (Fig. 3f). These findings were corroborated by phosphopeptide analysis. Dominant-negative PKBcaax abolished insulin-induced phosphorylation of peptides 1 and 2, but not 4. In contrast, RasV12 increased the phosphorylation of peptide 4, but RasN17 abolished insulin-induced phosphorylation of peptide 4, but not of 1 and 2 (Fig. 3g). Hence, Ras signalling through Ral is a strong candidate for mediating PKB-independent phosphorylation of AFX by insulin.

Members of the Forkhead family, to which AFX belongs, have been reported to bind to T(G/A)TTT motif-containing insulin-response elements (IREs), which inhibit the transcription of certain genes by insulin²¹. To investigate whether AFX can also bind such IREs, we fused the AFX DNA-binding domain (DB) carboxy terminal to GST, creating GST-AFX-DB. Purified GST-AFX-DB and GST-AFX could bind specifically to a radiolabelled oligonucleotide containing the IRE of the insulin-like growth factor-binding protein-1 (IGFBP-1) promoter, but not to an oligonucleotide containing a mutated IRE (Am2Bm2)²² defective in insulin responsiveness (Fig. 4a). To examine whether AFX can regulate transcription of the IGFBP-1 gene, we analysed the activity of a CAT reporter gene under the control of the IGFBP-1 promoter (1205-CAT)²². HA-AFX induced a pronounced increase in CAT activity (Fig. 4b). The effect of HA-AFX depended on an intact IRE, as the Am2Bm2 mutant blocked the effects of HA-AFX (Fig. 4b). Insulin treatment or co-expression of either gagPKB, RasV12 or Rlfcaax substantially inhibited the activity of the IGFBP-1 promoter (Fig. 4c). Transcriptional repression by insulin was regulated by the same pathways which induced AFX phosphorylation. Thus, only simultaneous blocking of Ras-mediated signalling to Ral (RasN17, RalN28) and PI(3)K-dependent signalling to PKB (wortmannin) could fully inhibit the insulin-induced repression of transcription (Fig. 4c). To determine whether phosphorylation of AFX by PKB was involved in the ability of AFX to regulate IGFBP-1 promoter activity, we examined the effect of a double mutant of HA-AFX lacking both PKB sites (HA-SASA). gagPKB had no effect on IGFBP-1 promoter activity when HA-SASA was expressed. In contrast, insulin was only slightly less able to inhibit expression of the CAT gene (Fig. 4c). This effect was probably due to insulin-induced Ras/Ral signalling, as RasV12 still inhibited transcription induced by HA-SASA, and blocking only Ras (RasN17) or Ral (RalN28), but not PI(3)K (wortmannin), prevented insulin from inhibiting AFX activity (Fig. 4c).

To investigate the underlying mechanism of insulin-induced inhibition of AFX, we constructed a fusion protein of full-length AFX with the Gal4 DNA-binding domain (Gal4-FL). Expression of Gal4-FL with a Gal4-luciferase reporter construct²³ increased luciferase-gene expression. Again, this transactivation was inhibited by insulin and involved both PKB and Ral (Fig. 4d). To determine which region of AFX was involved in the insulin-induced repression, we created Gal4 fusion proteins containing either the amino terminal or the C-terminal part of AFX (Gal4-N and Gal4-C, respectively). We found that the C-terminal part of AFX acted as a strong transcriptional transactivating domain. Transcription by Gal4-C, however, was not repressed by insulin (Fig. 4d). This, and the observation that insulin could repress activity of Gal4-FL, may indicate that the full-length AFX protein is required for inhibition by insulin.

Insulin induces the phosphorylation of AFX by way of both PI(3)K/PKB and a Ras/Ral signalling pathway. PKB phosphorylates

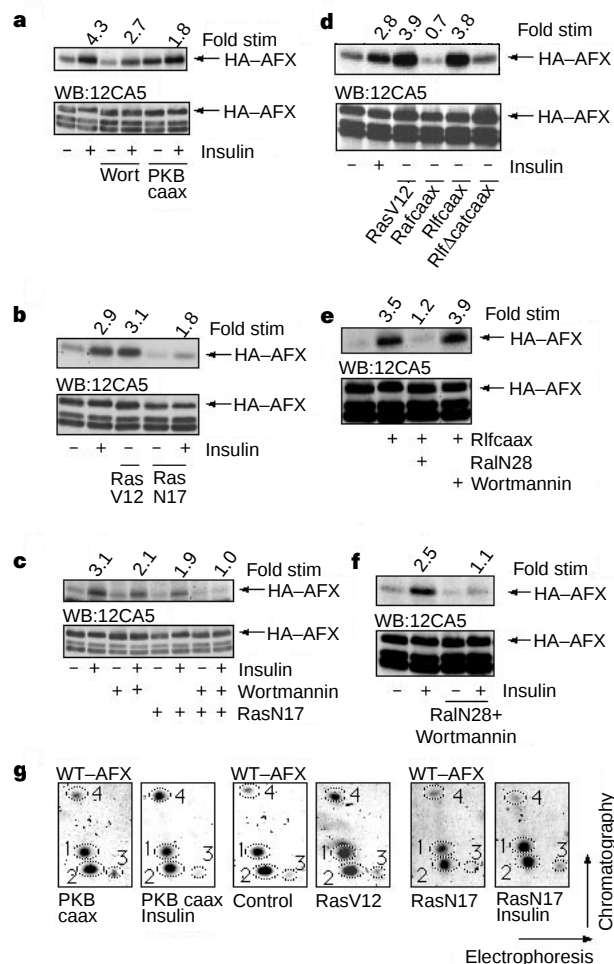


Figure 3 PKB cooperates with Ras in the phosphorylation of HA-AFX by insulin. **a-d, f**, A14 cells were left untreated (-) or treated with insulin for 30 minutes (+). Precipitates were analysed as in Fig. 1a. ³²P-labelled HA-AFX was isolated from transfected A14 cells co-expressing either PKBcaax or pretreated for 10 minutes with wortmannin (**a**), RasV12 or RasN17 (**b**), RasN17, pretreated for 10 minutes with wortmannin, or both (**c**), RasV12, Rlfcaax, Rlfcaax or RlfΔcatcaax (**d**), RalN28, pretreated for 10 min with wortmannin, or both (**f**). **e**, ³²P-labelled HA-AFX was isolated from transfected A14-cells co-expressing either Rlfcaax or Rlfcaax and RalN28 or co-expressing Rlfcaax and treated with wortmannin for 24 h. **g**, ³²P-labelled HA-AFX was isolated from transfected A14 cells co-expressing either PKBcaax, RasV12 or RasN17 and analysed as described in Fig. 2c.

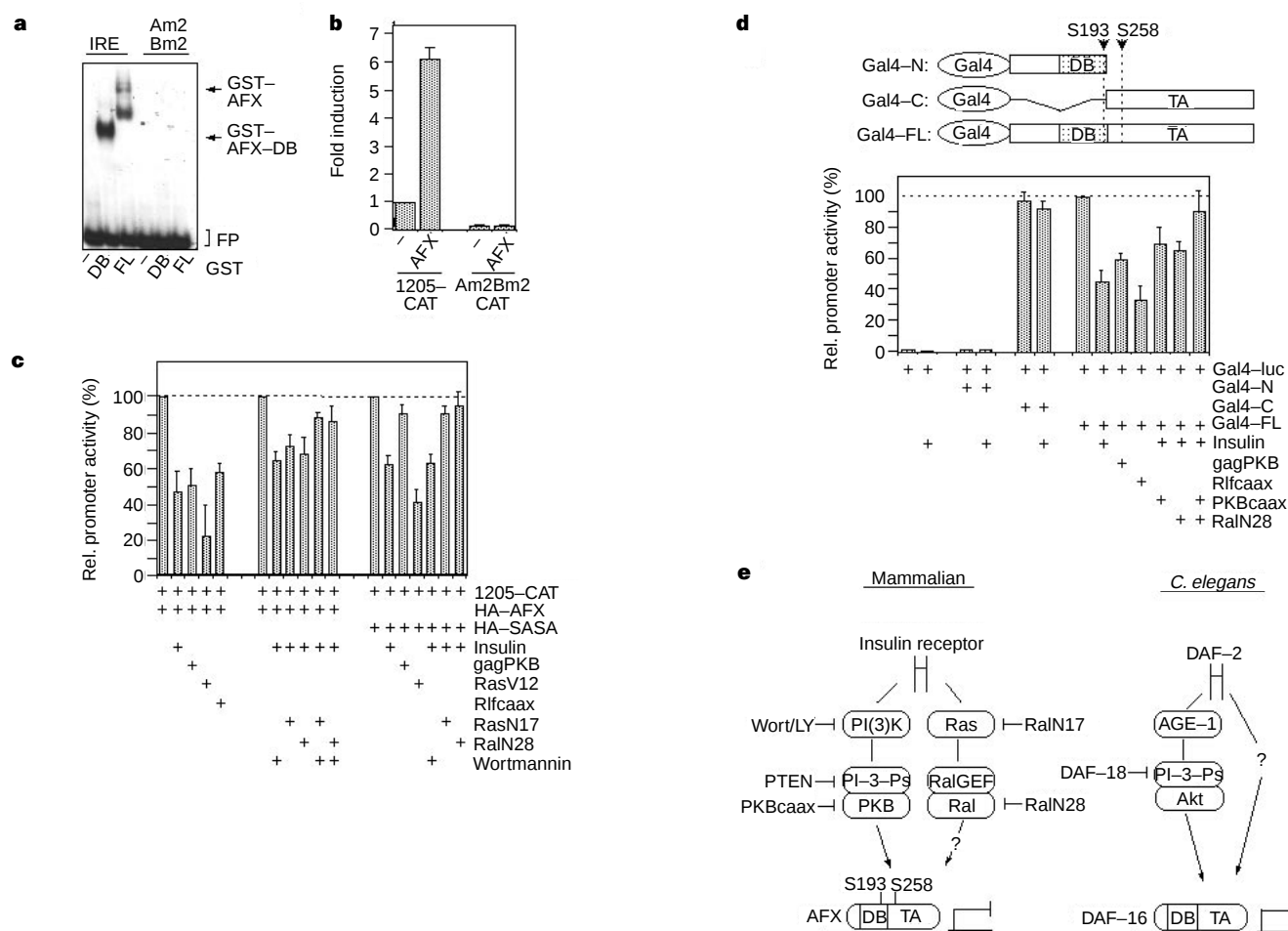


Figure 4 Phosphorylation of HA-AFX by PKB regulates transcriptional activity. **a**, GST protein alone (–), GST-AFX-DB (DB) or GST-AFX (FL) were incubated with ³²P-labelled oligonucleotides encompassing the IRE or a mutated IRE (Am2Bm2) of the IGF1R promoter. Faster migrating complexes in the GST-AFX lane represent GST-AFX break-down products. FP: free probe. **b**, A14 cells were transfected with the 1205-CAT reporter construct or the Am2Bm2-CAT reporter and co-transfected with (AFX) or without (–) wild-type HA-AFX. **c**, A14 cells were transiently transfected with the 1205-CAT reporter together with HA-AFX or HA-SASA either in the absence or presence of gagPKB, RasV12, Rlfcaax, RasN17 or RalN28. Treatment with insulin was for 16 h. Pretreatment with wortmannin was

for 10 min. **d**, A14 cells were transiently transfected with the Gal4-luc reporter construct together with either Gal4-N, Gal4-C or Gal4-FL in combination with either gagPKB, Rlfcaax, PKBcaax or RalN28. Treatment with insulin was for 7 h. **b–d**, Controls are set at 100%. Data are obtained from three independent experiments. **e**, A model for insulin-mediated inhibition of AFX-dependent transcription. Left, the insulin-induced pathways leading to phosphorylation and inactivation of AFX as described in the text. Inhibitors are indicated by blunted arrows. Right, pathways similar to those depicted in the left panel have been identified in *C. elegans*. DB: DNA-binding domain; TA: transactivation domain.

AFX predominantly on two residues, S193 and S258, and this phosphorylation by PKB can regulate the transcriptional activity of AFX (Fig. 4e). Although it is unlikely, we cannot formally exclude the possibility that, *in vivo*, a kinase with PKB-like specificity activated by PKB acts between PKB and AFX. Currently, no kinase acting downstream of Ral is known. Identification of the non-PKB target site on AFX will probably facilitate the isolation of such a kinase. Based on the use of a range of inhibitors and activators, MAP-kinase kinase (MEK), p90^{ras}, p70^{s6k}, protein-kinase A, protein-kinase C enzymes sensitive to 12-*D*-tetradecanoyl phorbol-13-acetate (TPA), p38/HOG1, glycogen synthase kinase-3 and calmodulin-kinase II do not seem to be involved in the phosphorylation of this PKB-independent target site (unpublished observations). Although activated PKB translocates to the nucleus, the physiological consequences of this translocation were unknown²⁴. Our data indicate that phosphorylation of the AFX transcription factor is one of the consequences of PKB relocalization. Our results confirm and extend the genetic data obtained in *C. elegans*. Complementation studies in *C. elegans* have suggested

that a PI(3)K (*age-1*)-independent route from the insulin receptor (*daf-2*) may lead to inactivation of AFX (*daf-16*), but the components of this route have yet to be identified (Fig. 4e and ref. 25). To our knowledge, a role for Ras in the regulation of *daf-16* in *C. elegans* has not been proposed. Our data identify the AFX transcription factor as an integration point for Ras and PI(3)K signalling, and as the first nuclear target of PI(3)K/PKB-dependent signalling. **Note added in proof.** Recently, Brunet *et al.* reported that the AFX-related transcription factor FKHR-L1 is also a substrate for PKB (Brunet, A. *et al.* Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. (*Cell* **96**, 857–868; 1999)). □

Methods

Cells and transfections. Insulin-receptor-overexpressing mouse NIH3T3 cells (A14) were grown as described¹. Insulin was added at 1 µg ml^{–1} and wortmannin (100 nM; Sigma) or LY294002 (10 µM; Sigma) was added 10 minutes before insulin. Transfections were carried out using the CaPO₄ method. Generally, 1 µg of the designated plasmid was added to 7 × 10⁵ cells.

Amounts of transfected complementary DNA were equalized using pBluescript KSII+.

Cloning and plasmids. The 5' region of the AFX gene was obtained by PCR of cDNA obtained from human leukaemic THP-1 cells. The 3' oligonucleotide (5'-AGCAGCTTGCTGCTGCTATCCATGGAG-3') contained the *NcoI* site located at position +578 of the AFX gene. The 5' oligonucleotide (5'-CACAGAAGGCCGTCGCGATCATAGAC-3') created an *NruI* site at position +26 of the AFX gene. Restriction sites are underlined. The 3' region of AFX was obtained from pUC-AFX1#3 (ref. 26). The *NruI/NcoI* fragment of the PCR product was ligated to the blunted *NcoI/HindIII* fragment of pUC-AFX1#3 and the resulting *NruI/HindIII* fragment was blunted and inserted into *SmaI*-cut pMT2-HA. pMT2-HA-AFX28A, S193A, S258A, S193A/S258A, S193T and S258T were generated by site-directed mutagenesis of the pMT2-HA-AFX cDNA using the following forward primers and subsequent complementary reverse primers: T28A (5'-CGCTCCTGCGCCTGGCCCC-3'), S193A (5'-GCCGGGCCGCCGCCATGGATA-3'), S258A (5'-CACGAAGCAGTGCAAA TGCCA-3'), S193T (5'-CCGCCGCCGGGCCGCCACCATGGATAGCAGCA GC-3'), and S258T (5'-CCGTCCACGAAGCAGTACAAATGCCAGCAGTG TC-3'). Base-pair changes are underlined. The double mutant was created by mutating the S193A cDNA using the S258A primers. pRP261-AFX-DBD (GST-AFX-DB) was generated by site-directed mutagenesis of pMT2-HA-AFX to create a *KpnI* site at position +249 and a *XbaI* site combined with an in-frame stop codon at position +623 of the AFX gene, using the following forward primers and subsequent complementary reverse primers: *KpnI* (5'-CGGAATCCTGGGGGCGGTACAGGTCCTCGGAAGGG-3') and *XbaI* (5'-GCCGCGAGTAAAGCCCTCTAGAAGAAACCATCTGTGC-3'). Restriction sites are underlined and the stop codon is in italic. Next, the *KpnI/XbaI* fragment was ligated into *KpnI/XbaI* cut pRP261. GST-AFX was created by ligating a *Sall/XhoI* fragment from pMT2-HA-AFX into *Sall* cut dephosphorylated pRP261. Gal4-N was created by ligating a blunted *Sall/NdeI* fragment of pMT2-HA-AFX into *XmaI* cut and blunted pSG424. Gal4-C was created by ligating a blunted *NcoI/EcoRI* fragment of pMT2-HA-AFX into *SmaI* cut pSG424. Gal4-FL was created by ligating a blunted *Sall/EcoRI* fragment of pMT2-HA-AFX into *XmaI* cut and blunted pSG424. All generated constructs were verified by automated sequencing.

The following plasmids have been described: pSG5-gag and pSG5-gagPKB¹, p110caax and p110-KDcaax²⁷. pcDNA3-myrPKB and pcDNA3-myrPKB-KD were obtained from D. Stokoe.

³²P-orthophosphate labelling. *In vivo* labelling of A14 cells was performed as described¹.

Immunoprecipitations and western blotting. Immunoprecipitations using the mouse monoclonal 12CA5 antibody and western blot analysis were performed under standard conditions and as described¹.

***In vitro* kinase assay.** *In vitro* kinase assay using active baculo-PKB was performed as described¹¹. The relative activity of 1 µl purified active PKB was calculated to be 1 pmol PO₄ per min, using a peptide containing S258 of AFX as a substrate.

Phospho-amino-acid analysis and tryptic peptide mapping. ³²P-orthophosphate labelled HA-AFX was immunoprecipitated from A14 cells, electrophoresed and immobilized on PVDF membrane (Immobilon). Protein was cut from the membrane and treated as described previously for phospho-amino-acid analysis²⁸ or peptide mapping²⁹. Sequence-grade trypsin was from Boehringer-Mannheim.

Electrophoretic mobility shift assay (EMSA). GST-AFX-DB and GST-AFX were purified using a standard GST-fusion-protein-purification protocol²². Bandshift assay was performed using a standard protocol²² and using the following oligonucleotides and subsequent complementary oligonucleotides: IGFBP-1 (5'-CACTAGCAAAACAACTATTGTAACAC-3'), Am2Bm2 (5'-CACTAGCAACCATGCCATGGTTGAACAC-3').

Chloramphenicol acetyl transferase assay. CAT assays were done as described³⁰.

Luciferase assay. Luciferase assays were done as described²³.

Received 16 February; accepted 15 March 1999.

- Burgering, B. M. Th. & Coffey, P. J. Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. *Nature* **376**, 599–602 (1995).
- Franke, T. F. *et al.* The protein kinase encoded by the *akt* proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase. *Cell* **81**, 727–736 (1995).
- Cichy, S. B. *et al.* Protein kinase B/Akt mediates effects of insulin on hepatic insulin-like growth factor-binding protein-1 gene expression through a conserved insulin response sequence. *J. Biol. Chem.* **273**, 6482–6487 (1998).
- Wang, D. & Sul, H. S. Insulin stimulation of the fatty acid synthase promoter is mediated by the phosphatidylinositol 3-kinase pathway. Involvement of protein kinase B/Akt. *J. Biol. Chem.* **273**, 25420–25426 (1998).
- Ogg, S. *et al.* The forkhead transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* **389**, 994–999 (1997).
- Lin, K., Dorman, J. B., Rodan, A. & Kenyon, C. *daf-16*: and HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. *Science* **278**, 1319–1322 (1997).
- Kimura, K. D., Tissenbaum, H. A., Liu, Y. & Ruvkun, G. *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* **277**, 942–946 (1997).
- Morris, J. Z., Tissenbaum, H. A. & Ruvkun, G. A. Phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in *Caenorhabditis elegans*. *Nature* **382**, 536–539 (1996).
- Borkhardt, A. *et al.* Cloning and characterization of AFX, the gene that fuses to MLL in acute leukaemias with a t(X;11)(q13;q23). *Oncogene* **14**, 195–202 (1997).
- Alessi, D. R., Caudwell, F. B., Andjelkovic, M., Hemmings, B. A. & Cohen, P. Molecular basis for the substrate specificity of protein kinase B; comparison with MAPKAP kinase-1 and p70 S6 kinase. *FEBS Lett.* **399**, 333–338 (1996).
- Van Weeren, P. C., de Bruyn, K. M., de Vries-Smits, A. M., van Lint, J. & Burgering, B. M. Essential role for protein kinase B (PKB) in insulin-induced glycogen synthase kinase 3 inactivation. *J. Biol. Chem.* **273**, 13150–13156 (1998).
- Cheatham, B. *et al.* Phosphatidylinositol 3-kinase activation is required for insulin stimulation of pp70 S6 kinase, DNA synthesis, and glucose transporter translocation. *Mol. Cell Biol.* **14**, 4902–4911 (1994).
- Pronk, G. J. & Bos, J. L. The role of p21ras in receptor tyrosine kinase signalling. *Biochim. Biophys. Acta* **1198**, 131–147 (1994).
- Medema, R. H., Wubbolts, R. & Bos, J. L. Two dominant inhibitory mutants of p21ras interfere with insulin-induced gene expression. *Mol. Cell Biol.* **11**, 5963–5967 (1991).
- Kauffmann-Zeh, A. *et al.* Suppression of c-Myc-induced apoptosis by Ras signalling through PI(3)K and PKB. *Nature* **385**, 544–548 (1997).
- Urano, T., Emkey, R. & Feig, L. A. Ras-GTPases mediate a distinct downstream signaling pathway from Ras that facilitates cellular transformation. *EMBO J.* **15**, 810–816 (1996).
- Burgering, B. M. & Bos, J. L. Regulation of Ras-mediated signalling: more than one way to skin a cat. *Trends Biochem. Sci.* **20**, 18–22 (1995).
- Wolthuis, R. M. F., de Ruiter, N. D., Cool, R. H. & Bos, J. L. Stimulation of gene induction and cell growth by the ras effector Rlf. *EMBO J.* **16**, 6748–6761 (1997).
- Leevers, S. J., Paterson, H. F. & Marshall, C. J. Requirement for Ras in Raf activation is overcome by targeting Raf to the plasma membrane. *Nature* **369**, 411–414 (1994).
- Wolthuis, R. M. F. *et al.* RasGDS-like factor (RIF) is a novel Ras and Rap1A-associating protein. *Oncogene* **13**, 353–362 (1996).
- Untermyer, T. G. *et al.* Hepatocyte nuclear factor-3 (HNF-3) binds to the insulin response sequence in the IGF binding protein-1 (IGFBP-1) promoter and enhances promoter function. *Biochem. Biophys. Res. Commun.* **203**, 1835–1841 (1994).
- O'Brien, R. M. *et al.* Hepatic nuclear factor-3 and hormone regulated expression of the PEPCK and IGFBP-1 genes. *Mol. Cell Biol.* **15**, 1747–1758 (1995).
- Duyndam, M. C., van Dam, H., van der Eb, A. J. & Zantema, A. The CR1 and CR3 domains of the adenovirus type 5 E1A proteins can independently mediate activation of ATF-2. *J. Virol.* **70**, 5852–5859 (1996).
- Andjelkovic, M. *et al.* Role of translocation in the activation and function of protein kinase B. *J. Biol. Chem.* **272**, 31515–31524 (1997).
- Paradis, S. & Ruvkun, G. *Caenorhabditis elegans* Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. *Genes Dev.* **12**, 2488–2498 (1998).
- Peters, U., Haberhausen, G., Kostrzewa, M., Nolte, D. & Müller, U. AFX1 and p54^{mb}: fine mapping, genomic structure, and exclusion as candidate genes of X-linked dystonia parkinsonism. *Hum. Genet.* **100**, 569–572 (1997).
- Didichenko, S. A., Tilton, B., Hemmings, B. A., Ballmer-Hofer, K. & Thelen, M. Constitutive activation of protein kinase B and phosphorylation of p47phox by a membrane-targeted phosphoinositide 3-kinase. *Curr. Biol.* **6**, 1271–1278 (1996).
- Boyle, W. J., Van der Geer, P. & Hunter, T. Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol.* **201**, 110–149 (1991).
- Luo, K. X., Hurley, T. R. & Sefton, B. M. Cyanogen bromide cleavage and proteolytic peptide mapping of proteins immobilized to membranes. *Methods Enzymol.* **201**, 149–152 (1991).
- De Groot, R. P. & Kruijer, W. Transcriptional activation by TGF beta 1 is mediated by the dyad symmetry element (DSE) and the TPA responsive element (TRE). *Biochem. Biophys. Res. Commun.* **168**, 1074–1081 (1990).

Acknowledgements. We thank K. Reedquist, F. Zwartkruis and P. van der Vliet for critically reading the manuscript and D. Nolte for the pUC-AFX1#3 plasmid. This work was supported by grants from NWO, KWF and CBG, and by the NIH (D.R.P.).

Correspondence and requests for materials should be addressed to B.M.Th.B. (e-mail: b.m.t.burgering@med.uu.nl).

Experimental values

A selection of new products for the lab with the emphasis – in an issue distributed at this year's FASEB gathering in Washington DC – on experimentation and the controlling the hazards it can bring.

ADC:OSI 1FL

From ADC BioScientific www.adc.co.uk

A compact portable chlorophyll fluorometer

The ADC:OSI 1FL is a hand-held chlorophyll fluorometer for the analysis of plant stress. The unit is battery operated and weighs only 2 kg. Fluorescence parameters are displayed on an LCD screen. Programming is via a simple keypad and automatic test modes can be set up so that repetitive experiments can be conducted by a single key operation.

Reader Service No. 100

The GasHound 800

From LI-COR www.licor.com

A precision CO₂ analyser

The GasHound 800 is a single-path, dual-wavelength, non-dispersive infrared gas analyser. It is easily portable and can also operate from the mains, batteries or most DC power supplies. Recommended applications include industrial process control and agricultural research.

Reader Service No. 101

Centrifuge trolleys

From Kendro www.kendro.com

Ergonomic safety trolleys for table-top centrifuges

These new base tables are available for both refrigerated and non refrigerated Heraeus centrifuges. Tables are available with various height and weight-bearing capacities to suit different applications. Working height is an ergonomic 90 to 100 cm. For safety, drawers are designed so that they remain tightly



A good runner: this draw won't open on the move.

closed if the trolley is moved. The trolleys are of reinforced metal construction, so they remain vibration free and minimize noise during a centrifuge run.

Reader Service No. 102

HMI 41

From Vaisala www.vaisala.com

A hand-held humidity/dewpoint indicator

As well as displaying humidity, dewpoint and temperature readings, the HMI 41 can calculate wet bulb temperature, absolute humidity and mixing ratio signalling. It measures humidity from 0 to 100 %RH at temperatures from -40 to +100 °C and will operate at temperatures of up to +180 °C for short periods of time. Temperature readout is available in either degrees Centigrade or Fahrenheit and calculated variables are available in metric or non-metric units, all displayed on a built-in LCD screen. A range of probes is available for specialist applications.

Reader Service No. 103

MSDS

From Packard Instrument Company

www.packardinst.com

Material Safety Data Sheets are available on Packard's web site

The list of MSDSs accessible on Packard's web site include all of the most popular liquid scintillation cocktails and reagents. The MSDSs provide comprehensive information on all aspects of health and safety with respect to liquid scintillation cocktails and reagents, and include composition, hazard, first aid, fire fighting, handling, exposure, toxicological, ecological, disposal, transport and regulatory sections within the document.

Reader Service No. 104

Biohazardous Waste Container

From Nalge Nunc International

www.nalgenunc.com

A new small size for a biohazard waste containers

Nalge Nunc has introduced a compact 1.5 gallon (5.5 litre) version of the Nalgene autoclavable plastic waste containers for secondary containment of biohazardous waste materials. They permit convenient sterilization of waste, with no need to remove the autoclaving bag before decontamination. The port opening in the cover allows simple, one-handed disposal of



Bin thinking: now in a smaller size.

materials, while limiting exposure to infectious waste. Features include leak-proof side and bottom construction which will not rust, dent or puncture. The new durable polypropylene copolymer container holds a 12 × 24 in. (305 × 610 mm) autoclavable bag and complies with US OSHA Standard 29 CFR Part 1910.1030 for use as protection against blood-borne pathogens. The 5- and 15-gallon (19- and 57-litre) versions of the container remain in the range.

Reader Service No. 105

Protector XL

From Labconco

www.labconco.com

Fume hoods offering larger sizes and versatility

Protector XL Hoods are available in widths of 8, 10, 12 and 16 feet in both bench-top and walk-in styles. The standard internal depth is 30 inches, but 24 and 36 inch depths are also available. The chemical-resistant liner panels are available in several choices of material and the exterior is epoxy-coated steel. Internal lighting can be fluorescent or explosion-proof incandescent.

Reader Service No. 106

Chemical Storage

From Labcaire Systems

www.labcaire.co.uk

A new cabinet to eliminates fumes from the storage of pathology specimens

Labcaire's new Chemical Storage Cabinet (CSC) is designed to protect staff and eliminate fumes and odour associated with the storage of formalin-fixed pathology samples. The cabinet works by generating a continuous airflow into the cabinet, circulating it around all shelves and finally putting

through a fume removing activated carbon filter. The cabinet has fully adjustable shelves and a filtration system situated at the base of the cabinet for convenient filter changes. As a free standing system, the unit requires no connection to dusting or additional ventilation devices, thus ensuring both easy positioning and installation.

Reader Service No. 107

Clear-spin

From Seward

www.seward.co.uk

Safer centrifugation of hazardous biological and radioactive materials

The new Clear-spin microcentrifuge, launched at Pittcon this year, has been designed to allow easier and safer manipulation of hazardous samples material during centrifugation. The main safety feature is the new rotor disc, which also doubles as a storage rack. Sample preparation and manipulation can be performed with the tubes held in the rotor disc, prior to and after transfer to the rotor case. The rotor case is hermetically sealed and constructed from transparent polysulphone so that spillage or tube failure can be detected. Clear-spin incorporates a strobe to allow direct observation of separating components.

Reader Service No. 108

5415 D centrifuge

From Eppendorf

www.eppendorf.com

Latest version of an established bench-top centrifuge

As a successor to their 5415 C, Eppendorf has developed the Microcentrifuge 5415 D. Either run-time or rotational speed are entered into the dials as appropriate, and a digital display shows either the rotational speed or the relative centrifugal force as required. An aerosol-tight lid is available for the 24-position standard rotor for 1.5 ml



Another D-type with a high top speed.

and 2 ml test tubes. The 5415 D complies with safety regulations for the handling of infectious materials. With the aid of adapters, 0.2 ml, 0.4 ml, 0.5 ml and 0.6 ml test tubes can also be centrifuged. The centrifuge can accelerate to its maximum (16,110g) in less than 13 seconds.

Reader Service No. 109

GR 20.22

From Jouan

www.jouaninc.com

Avoid that walk to the centrifuge room

Jouan's claim for the GR 20.22 is that it is the industry's only high-speed centrifuge that can realistically be called "personal". Its small footprint allows this centrifuge to be stored and used in your own lab — not in a separate centrifuge room like traditional high-speed machines. The GR 20.22 can spin six 290-ml samples at 14,126 g and eighteen 16-ml samples at up to 42,037g, for less than half the cost of traditional high-speed machines.

Reader Service No. 110

The Free Exploratory Test

From Neurofit SA

www.tpgnet.net/neurofit

A new behavioural procedure for testing potential antineophobic drugs in mice

Neurofit is a company specializing in

ADVERTISEMENTS

NEW ANTIBODIES FOR UBIQUITIN AND PROTEASOME RESEARCH FROM AFFINITI

20S Proteasome α - & β -subunits: HC2, HC3, HC8, HC9, ι , XAPC7, zeta, HC10, HN3, Z, MECL1, LMP2, LMP7; 19S regulator ATPase & non-ATPase subunits; 11S regulator subunits: PA28 α , β and γ ; **Signalosome** subunits; PGP9.5; E2; Ub-conjugates

For data sheets or more information on AFFINITI's extensive range of products for ubiquitin and proteasome research visit our website at:

[HTTP://WWW.PROTEASOME.COM](http://www.proteasome.com)

AFFINITI Research Products Ltd., Mamhead Castle, EXETER, EX6 8HD, UK
Tel. +44/0 1626 891010; Fax: +44/0 1626 891090; URL: <http://www.affiniti-res.com>

The Mouse Kit™ for Gene Targeting in ES Cells

- ✓ Complete with embryonic stem (ES) cells, feeder cells and essential reagents (including FBS) for engineering one knockout mouse strain
- ✓ Low-passage AB2.2-Prime ES Cells* (129/SvEv)
- ✓ Uniform, high-quality ESQ feeder cells (mitotically inactivated STO cells) engineered for *neo* and *puro* resistance
- ✓ Prequalification of ESQ feeder cells and reagents for use with AB2.2-Prime ES Cells

*Manufactured by Lexicon Genetics Incorporated

Fax: (619) 535-0871

Tel: (619) 535-5400 or (800) 424-5444

www.stratagene.com

READER ENQUIRY NO. 38

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For more information please contact Dr. Joe Gray:
Phone/Fax: +44 191 222 8612

Email: JOE.GRAY@newcastle.ac.uk

*DNA sequencing (ABI 377-96KL, 877 Catalyst)
also available at competitive prices*

READER ENQUIRY NO. 5



BMA BIOMEDICALS AG

Your source for unique antibodies:

MCP-1
Scavenger Receptor CD163
Endothelial and Macrophage markers
Murine myeloid cell markers
hMRP8/14 (Calprotectin) ELISA
hIL-1 receptor antagonist ELISA

Call for your free catalogue:
++41 61 811 6222 (phone)
++41 61 811 6006 (fax)
info@bma.ch

READER ENQUIRY NO. 1

READER ENQUIRY NO. 6

Diagnostic Reagents

Pefachrome®-Series: chromogenic and fluorogenic substrates

Pefabloc®-Series:
serine proteinase inhibitors

Purified snake venom proteins
(e.g. Protac®, Ecarin, Batroxobin)

PENTAPHARM LTD

CH 4002 Basel Fax: +41 61 319 96 19
<http://www.pentapharm.ch>

READER ENQUIRY NO. 16

neuroscience preclinical research, performing contract research on animal models for various neurological or psychiatric disorders. A new behavioural procedure is now available for the evaluation of emotional responses of animals towards a new environment and reveals their endogenous anxiety, 'trait anxiety'. In this paradigm, Balb/c mice have display neophobic reactions. The paradigm, the free exploratory test, is also suitable for evaluation of sedative or psychostimulant effects of substances.

Reader Service No. 111

Respiromax

From Columbus Instruments

www.colinst.com

Tidal volume respiratory function computer

This new system accommodates 1 to 8 animals for precision measurements of the tidal volume, respiration rate, maximum expiratory flow, inspiration time, expiration time and tidal volume curve during exposure to toxins and irritants. The system includes

animal holders with designed to allow an unobtrusive but secure seal around the animal's neck separating the animal's body from the head breathing in the inhalation exposure chamber. Respiromax is controlled by a PC. During the experiment, the user has the ability to observe the tidal volume curve and numerical values of respiratory parameters on computer screen or store snapshots of the curve and computed parameters to the hard disk.

Reader Service No. 112

Seripettor

From Brand

www.brand.de

A novel low-cost bottle-top dispenser

Seripettor offers a realistic alternative to high-end dispensers. The materials used for construction mean it can be used for dispensing many common laboratory liquids: dilute acids, buffers, biological solutions, polar solvents, and strong alkaline solutions. The modular design makes cleaning and maintenance easy. The Seripettor allows semiautomatic dispensing, as the



Seripettors: but do they work with Cointreau?

spring-loaded piston refills the cylinder automatically after dispensing. It comes in two sizes: 1–10 ml and 2.5–25 ml.

Reader Service No. 113

[3H]-SNC 121 and [3H]-HU 243

From Semat Technical

www.semat.co.uk

Receptor radioligands

Two new TOCRIS compounds are now available from Semat Technical. [³H]SNC 121, the tritium saturated analogue of SNC-80, is a potent and selective non-peptide δ -opioid receptor ligand, which together with its non-radiolabelled counterpart, is an important pharmacological tool for studying the δ -opioid. [³H]SNC 121 labels a single class of δ -opioid receptor with high affinity and selectivity (>1,000-fold). [³H]HU 243 is the tritiated form of the potent cannabinoid receptor agonist HU 243, binding to cannabinoid receptors with a K_d of 45 pM in rat synaptosomal membranes.

Reader Service No. 114

BetaPette

From Continental Laboratory Products

www.conlab.com

Continuously adjustable micropipettes

CLP's Beta Pette series of air displacement pipettors are low-maintenance pipettors capable of delivering aliquots from 0.2 μ l to 5 ml. They feature light-weight polymer handles and removable/adjustable tip ejectors and are fully autoclavable. CLP also offers an in-house calibration and maintenance programme. Address for e-mail enquiries is mail@conlab.com

Reader Service No. 115

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENT

Check out our homepage at <http://www.resgen.com>

Peptides
\$14/residue
10 residue minimum

Polyclonal Antibodies - \$1,195

Please contact us about complete services for your custom peptide and polyclonal antibody needs.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

READER ENQUIRY NO. 4

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm!

READER ENQUIRY NO. 54

The Bedside Book: Genius and Eccentricity in Science 1869–1953

Edited by Walter Gratzer, with a foreword by Stephen Jay Gould

Price: £19.95/US\$29.95 + postage
Discount rate available for current subscribers.
Contact: Nature Marketing
Tel. +44 (0) 171 843 4985
Fax. +44 (0) 171 843 4998
E-mail: subscriptions@nature.com

Detect residual DNA with Wako's DNA Extractor Kit or Isolate Genomic DNA with DNA Extractor WB Kit

Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 39